

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LE PROCESSUS DE DÉCODAGE ET DE PRODUCTION DE MOTS
DÉRIVÉS CHEZ DES APPRENANTS ADULTES
DU FRANÇAIS LANGUE SECONDE

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN LINGUISTIQUE

PAR
ISABELLE LÉPINE

SEPTEMBRE 2007

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Cette recherche a été rendue possible grâce au concours de très nombreuses personnes et ce, tout au long des différentes étapes qui m'ont amené à terminer ce mémoire.

J'aimerais d'abord remercier la responsable des cours de français des niveaux avancés de l'École de langues, Mme Marie-Cécile Guillot, ainsi que les professeurs Francine Campeau et Isabelle Gingras de m'avoir permis de réaliser ma collecte de données dans leur classe de français langue seconde et ce, pendant les heures de cours. Je tiens également à remercier tout particulièrement les étudiants de ces classes qui ont non seulement participé avec enthousiasme à cette recherche, mais qui m'ont de plus encouragé dans ma démarche.

J'aimerais également remercier Mme Marie Labelle, professeure au département de linguistique de l'UQAM, pour ses conseils lors du codage de mes données. Un merci tout particulier à ma directrice de mémoire, Mme Lori Morris, pour son aide si précieuse, sa très grande souplesse, ses réflexions pertinentes ainsi que son réel enthousiasme envers un sujet comme le mien.

Je ne saurais passer sous silence toute la collaboration et le soutien moral de mon conjoint, Stéphane Arseneau, qui a toujours été là pour moi et qui m'a permis de finaliser ce mémoire en s'occupant de notre petit garçon, me laissant ainsi quelques heures de répit pour écrire.

Finalement, je voudrais remercier mon petit garçon Jérémy pour avoir toujours été aussi compréhensif envers sa mère qui devait passer de longues heures devant son écran d'ordinateur. Merci pour ta patience et tes sourires coquins.

TABLE DES MATIÈRES

LISTE DES FIGURES	V
LISTE DES TABLEAUX.....	VI
RÉSUMÉ	VIII
INTRODUCTION	1
CADRE THÉORIQUE ET PROBLÉMATIQUE	4
1.1 La morphologie dérivationnelle.....	5
1.1.1 La morphologie dérivationnelle en langue première	6
1.1.2 La morphologie dérivationnelle en langue seconde	8
1.1.3 Modèles théoriques du traitement de la morphologie en L2.....	12
1.1.3.1 The Bilingual Interactive Activation Plus Model (BIA+)	13
1.1.4 Influence de la L1 dans le traitement des mots dérivés de la L2.....	17
1.1.4.1 Le modèle BIA+ et le rôle de la langue première	17
1.1.4.2 Recherches sur l'influence de la L1 dans le traitement de la morphologie en L2	20
1.2 Les stratégies d'apprentissage	23
1.2.1 Taxinomies	23
1.2.2 Les stratégies d'apprentissage et la morphologie dérivationnelle.....	27
1.2.3 Recherches sur l'influence de la L1 sur les stratégies d'apprentissage ...	28
MÉTHODOLOGIE	32
2.1 Présentation de la recherche	32
2.1.1 Questions de recherche	33
2.1.2 Hypothèses de recherche.....	35
2.2 Participants	35
2.3 Instruments de mesure	37
2.3.1 Le C-test.....	38
2.3.1.1 Élaboration et contenu des C-tests.....	41
2.3.2 La tâche de rappel stimulé.....	42
2.3.2.1 Élaboration et contenu de la tâche de rappel stimulé	45
2.4 Déroulement et collecte des données.....	45
2.5 Codage et analyse des données.....	48

2.5.1 Codage – C-tests	48
2.5.2 Codage - Tâche de rappel stimulé	49
2.5.3 Analyses	51
PRÉSENTATION ET DISCUSSION DES RÉSULTATS	54
3.1 Résultats – C-tests.....	54
3.1.1 Résultats – C-test 1 et C-test 2	55
3.1.1.1 Répartition des participants.....	55
3.1.1.2 Problèmes encourus lors du codage des C-tests.....	58
3.1.1.3 Résultats - Réussite globale.....	61
3.1.1.4 Résultats - Réussite des mots dérivés	64
3.1.1.5 Résultats – Mots discriminants.....	66
3.1.2 Résultats – C-test 3	77
3.1.2.1 Répartition des participants.....	78
3.1.2.2 Résultats – Réussite globale.....	79
3.1.2.3 Résultats – Réussite des mots dérivés.....	80
3.1.2.4 Réussite – Mots discriminants.....	82
3.2 Résultats – Tâche de rappel stimulé.....	86
3.2.1 Répartition des participants.....	87
3.2.2 Problèmes encourus	87
3.2.3 Résultats - Tâche de rappel stimulé.....	89
3.2.3.1 Résultats – Catégories de stratégies les plus nommées	90
3.2.3.2 Stratégies discriminant les deux groupes	93
3.2.3.3 Correspondance entre stratégies et réponses au 3 ^e C-test.....	100
3.3 Synthèse	104
3.3.1 Résumé – Résultats des C-tests	104
3.3.2 Résumé - Résultats de la tâche de rappel stimulé.....	105
3.3.3 Synthèse – Hypothèses de recherche	106
CONCLUSION	109
APPENDICE A : C-TESTS	114
APPENDICE B : PROTOCOLE DE RECHERCHE.....	124
APPENDICE C : PROJET DE RECHERCHE.....	128
RÉFÉRENCES.....	131

LISTE DES FIGURES

Figure	Page
1.1 The Bilingual Interactive Activation Plus Model (Tiré de Dijkstra et Van Heuven, 2002)	14
3.1 Réussite globale – C-test 3 – LPA \ LBA	79
3.2 Réussite des mots dérivés – C-test 3 – LPA \ LBA.....	81
3.3 Catégories de stratégies les plus identifiées et celles les moins nommées	91
3.4 Stratégies discriminantes	94
3.5 Correspondance entre les stratégies identifiées et les réponses au troisième C-test	102

LISTE DES TABLEAUX

Tableau	Page
1.1 Taxinomie des indices servant à décoder des mots nouveaux dans la L2. (Tiré de Haastrup, 1987)	25
1.2 Learning strategies : Self-reports (Tiré de Gu et Johnson, 1996)	26
2.1 Profil des participants selon le sexe	36
2.2 Profil des participants selon la L1	36
2.3 Grille de codage des C-tests	48
2.4 Grille de codage des stratégies	50
2.5 Grille de codage : correspondance entre stratégie et réponse	53
3.1 Nombre de C-tests rejetés selon le type d'explication	56
3.2 Répartition des participants selon le sexe et la L1 C-tests 1 et 2	57
3.3 Réussite globale selon le sexe – C-tests 1 et 2	62
3.4 Réussite globale selon les caractéristiques morphologiques du système d'écriture de la L1 - C-tests 1 et 2	63
3.5 Réussite des mots dérivés selon les caractéristiques morphologiques du système d'écriture de la L1 – C-tests 1 et 2 .	65
3.6 Mots dérivés les plus réussis – C-test 1	67
3.7 Mots dérivés les moins réussis – C-test 1	70
3.8 Mots dérivés les plus réussis – C-test 2	72
3.9 Mots dérivés les moins réussis – C-test 2	74
3.10 Mots discriminants – C-tests 1 et 2	76
3.11 Répartition des participants selon leur L1 – C-test 3	78

3.12	Mots dérivés les plus réussis – C-test 3	82
3.13	Mots dérivés les moins réussis – C-test 3	84
3.14	Mots discriminants – C-test 3	86
3.15	Codage de la correspondance entre stratégie et réponse	100

RÉSUMÉ

À Montréal, nombreux sont les nouveaux immigrants qui choisissent d'apprendre le français langue seconde. Cette situation pose de nouveaux défis aux enseignants qui doivent composer avec une clientèle de plus en plus diversifiée culturellement et linguistiquement.

Par cette recherche, nous avons voulu vérifier si la langue maternelle des apprenants adultes avait un impact sur leur capacité à traiter les mots dérivés en français ainsi que sur les stratégies qu'ils utilisent pour décoder et produire ces mots dérivés. Pour vérifier nos hypothèses, deux groupes d'apprenants de langues maternelles différentes ont d'abord été testés à l'aide de C-tests et ont ensuite participé à une tâche de rappel stimulé.

Les résultats montrent que les caractéristiques morphologiques du système d'écriture de la langue maternelle affecte d'une manière significative la capacité à reconnaître et à produire les mots dérivés sans influencer toutefois les stratégies déployées pour traiter ces mots. Les différences significatives observées ne seraient pas causées par un manque de connaissances de la morphologie dérivationnelle du français chez certains sujets, mais bien plutôt par un manque de connaissances lexicales. Cette dernière hypothèse reste toutefois à confirmer.

INTRODUCTION

Le Canada a toujours été une destination de choix pour les immigrants et le Québec, depuis quelques années, connaît un véritable engouement¹. Une des conséquences de cette hausse de l'immigration est une augmentation importante du nombre de cours de français langue seconde dans les grandes villes. À Montréal, contexte de notre recherche, les écoles de langues et autres programmes de français pour les immigrants adultes se sont multipliés en l'espace de quelques années : cours de français de base offerts par le ministère de l'Immigration et des Communautés culturelles, cours du soir offerts dans des écoles de langues de la région, cours de français offerts dans les universités et les centres communautaires, et ainsi de suite.

Malgré toute cette diversité, la seule et même priorité demeure : celle de réussir à offrir un enseignement de qualité adapté à ces étudiants qui doivent, le plus rapidement possible, s'intégrer linguistiquement et culturellement dans la société québécoise. Pour atteindre cet objectif, il est essentiel de comprendre comment ces adultes, certains très scolarisés et d'autres peu scolarisés, réussissent à maîtriser le français. Quelles sortes de difficultés éprouvent-ils ? Trébuchent-ils tous sur les mêmes aspects ? Ont-ils tous les mêmes forces ? Si oui, quelles sont ces forces ? Si non, qu'est-ce qui différencie ces étudiants ? Toutes ces questions nous démontrent que, plus que jamais, la recherche sur l'acquisition du français langue seconde a sa place dans les milieux d'enseignement.

Jusqu'à maintenant, l'acquisition de plusieurs catégories grammaticales du français a fait l'objet d'études en langue seconde : le genre (Holmes et De la Bâtie, 1999), les pronoms (Duffield, White, Bruhn de Garavito, Montrul et Prevost, 2002), les temps de verbes (Paprocka, 2002), etc. Côté morphologie, il

¹ Selon des données récentes du ministère de l'Immigration et des Communautés culturelles, le Québec a accueilli 43 373 immigrants en 2005, alors qu'il n'en avait admis que 32 440 en 2000.

est intéressant de constater que l'intérêt des chercheurs a surtout porté sur l'acquisition de la morphologie flexionnelle (Prevost, 2004; Bourdin, Leuwers, Bourbon, Chevrot, Schelstraete, 2004), la morphologie dérivationnelle ayant été en quelque sorte délaissée. Cette absence d'intérêt pour la morphologie dérivationnelle se constate également dans le milieu scolaire où elle est un sujet peu ou pas traité dans les manuels et dans les cours de français langue seconde offerts aux apprenants adultes au Québec. Ces deux constatations sont étonnantes étant donné qu'en français, plus de sept mots sur dix sont morphologiquement complexes (Rey-Debove, 1984). Par conséquent, de deux choses l'une : soit les apprenants n'ont pas besoin d'une formation explicite en morphologie dérivationnelle car ils acquièrent le nécessaire implicitement, soit l'enseignement du vocabulaire ne répond pas aux besoins des apprenants. À notre avis, il est essentiel de comprendre comment les apprenants adultes du français langue seconde font pour reconnaître, analyser et mémoriser les mots dérivés. Non seulement le bagage lexical de l'apprenant en dépend, mais également son habileté à comprendre textes et conversations.

En vue d'éclairer la question du traitement de la morphologie dérivationnelle en français langue seconde, nous avons formulé les deux questions de recherche suivantes :

- a) Est-ce que tous les apprenants adultes et ce, peu importe leur bagage linguistique, réussissent avec la même facilité une tâche de décodage et de production de mots suffixés en français langue seconde ?
- b) Est-ce que ces mêmes apprenants utilisent des stratégies similaires pour décoder et produire des mots dérivés ?

Afin de nous aider à répondre à ces deux questions, 63 apprenants adultes du français langue seconde ont participé à notre recherche en complétant deux C-tests qui ont été administrés en classe à sept semaines d'intervalle. Cette première série d'épreuves nous a permis de mesurer en pourcentage (%) leur performance au

niveau du décodage et de la production de mots dérivés. De ce nombre total de participants, vingt-quatre d'entre eux ont poursuivi l'expérience en complétant un troisième C-test, sept semaines plus tard, et en acceptant de répondre à des questions lors d'une tâche de rappel stimulé, nous permettant ainsi d'identifier les stratégies qui avaient été utilisées lors du C-test.

Nos deux instruments d'évaluation, le C-test et la tâche de rappel stimulé, ont été conçus spécifiquement pour une population d'apprenants adultes du français langue seconde. Mentionnons, pour terminer, que notre échantillon provenait du milieu universitaire et était composé d'apprenants aux profils linguistiques très variés.

Ce mémoire présente donc l'état de la question, la problématique, la méthodologie et les résultats de notre recherche. Dans le premier chapitre, nous donnons une brève description de la morphologie dérivationnelle avant de présenter les différentes recherches en langue seconde qui ont porté sur le sujet. Nous concluons en abordant le thème des stratégies. Dans le deuxième chapitre, nous présentons nos hypothèses de recherche ainsi que les démarches que nous avons entreprises afin de valider (ou infirmer) ces hypothèses. Nous discutons ensuite de nos participants, du choix et de l'élaboration de nos instruments de mesure, de notre collecte de données ainsi que du codage et de l'analyse des résultats recueillis. Dans le troisième chapitre, il est question des résultats obtenus et de leur signification en ce qui a trait à notre sujet. Ce qui nous amène ensuite à la conclusion de ce mémoire, conclusion dans laquelle nous abordons les limites de cette recherche ainsi que les implications pédagogiques souhaitables.

CHAPITRE I

CADRE THÉORIQUE ET PROBLÉMATIQUE

Lorsque nous apprenons une nouvelle langue, notre premier contact avec celle-ci commence bien souvent par l'apprentissage de mots de vocabulaire. En effet, le lexique est une partie importante de l'acquisition d'une langue (Laufer, 1997). Mais que veut réellement dire « apprendre » un mot de vocabulaire ? Qu'est-ce que cela implique ? De nombreux chercheurs se sont penchés sur la question (Nation, 1990; Gass et Selinker, 2001; Chapelle, 1994, etc.) et grâce à eux, aujourd'hui, on sait que connaître un mot de vocabulaire signifie maîtriser plusieurs aspects de ce mot dont, entre autres, sa structure interne, comme l'exprime si bien Laufer (1997) : « In order to know a word, it is necessary to know (...) word structure – the basic free morpheme (or bound root morpheme) and the common derivations of the word and its inflections. » (p. 141).

La morphologie dérivationnelle fait donc partie, en quelque sorte, des savoirs importants que doit acquérir tout apprenant qui souhaite maîtriser les mots de vocabulaire d'une nouvelle langue. Dans la première partie de ce chapitre, nous discuterons de la morphologie dérivationnelle en la définissant, tout d'abord, et en présentant ensuite l'état des connaissances que nous avons actuellement sur son acquisition et son traitement en langue première (L1) et en langue seconde (L2). Nous aborderons par la suite les modèles théoriques qui se sont intéressés au fonctionnement du lexique mental avant de terminer par une réflexion sur l'influence de la L1 dans le traitement de la morphologie en L2.

Dans la deuxième partie de ce chapitre, nous traiterons des stratégies d'apprentissage. Nous définirons d'abord ce qu'est une stratégie d'apprentissage et nous présenterons ensuite quelques taxinomies. Par la suite, nous expliquerons les stratégies d'apprentissage en lien avec la morphologie et nous terminerons en

présentant les résultats de quelques recherches qui se sont intéressées à l'influence de la L1 sur les stratégies utilisées en L2.

1.1 La morphologie dérivationnelle

La morphologie en général s'intéresse à la structure des mots, à leur formation ainsi qu'à la variation de leurs formes (Brousseau et Nikiema, 2001). En français, on distingue deux types de morphologie : la morphologie lexicale, qui étudie les procédés qui régissent la formation des mots, et la morphologie flexionnelle « qui traite des variations de la forme des mots selon les catégories du nombre, du genre, de la personne, etc. » (Riegel, 1994). Pour les besoins de notre recherche, nous nous limiterons à la morphologie lexicale.

En français, il y a deux mécanismes qui permettent la construction des mots : la dérivation et la composition. La dérivation consiste en l'ajout d'affixes (préfixe et suffixe) à une unité lexicale, ou mot simple, que l'on appelle « base ». Quant à la composition, il s'agit d'un procédé qui consiste en la combinaison de deux mots simples ou plus, comme *porte-monnaie* par exemple. Dans le cadre de notre étude, nous nous intéresserons à la dérivation des mots à l'aide d'affixes, et plus spécifiquement à la connaissance des mots suffixés. Plusieurs raisons nous ont amené à faire ce choix. Tout d'abord, les suffixes sont nombreux et ils n'ont pas tous le même degré de productivité. Cette variété est un aspect intéressant quand on veut mesurer les connaissances morphologiques. Ensuite, l'ajout de suffixes est régi par des règles bien précises (ex. : le suffixe *-ment* s'ajoute à la base d'un adjectif ou à la base d'un verbe pour former un nom) et il devient alors très facile de repérer ceux qui commettent des erreurs imputées au manque de connaissances de ces règles. Finalement, ce qui est intéressant avec les suffixes, c'est qu'ils nous permettent d'avoir une bonne idée du bagage lexical de l'apprenant. Puisque l'ajout d'un suffixe à une base en change généralement la catégorie

grammaticale², l'apprenant qui doit manipuler des suffixes doit ainsi démontrer qu'il connaît, entre autres, ce qu'est une famille de mots.

1.1.1 La morphologie dérivationnelle en langue première

La morphologie dérivationnelle a longtemps été mise à l'écart comme source potentielle d'informations sur le développement du langage (Ravid, 2004). Mais depuis quelques années, des chercheurs clés, tels Joanne F. Carlisle et Diana Mahony, s'y sont intéressés. L'une des premières questions posées concerne la période d'acquisition de la morphologie dérivationnelle : quand les enfants développent-ils une connaissance implicite de la structure interne des mots ? Bien que la structure de la langue soit apprise à l'oral en premier, il semblerait que les connaissances morphologiques se développeraient avec l'apprentissage de l'écriture, et par le fait même de la lecture. Certains chercheurs proposent les premières années du primaire, puisque ce sont durant ces premières années que commence à apparaître un lien entre la conscience morphologique et la lecture (Carlisle et Nomanbhoy, 1993; Casalis et Louis-Alexandre, 2000; Mahony, Singson et Mann, 2000). D'autres suggèrent les années pivots entre le primaire et le secondaire (Leong, 1989; Tyler et Nagy, 1989) ou encore la période s'échelonnant du secondaire au collégial (Mahony, 1994).

Si la période exacte de l'acquisition de la morphologie dérivationnelle est encore un sujet de débat, il n'en demeure pas moins que c'est durant la période du « Later language development » que l'on constate une croissance exponentielle du vocabulaire et, par le fait même, des habiletés morphologiques (Clark, 1993; Nagy et Anderson, 1984; Nagy et Herman, 1987; Nippold, 1998). Toutefois, même à ce stade, les habiletés morphologiques ne seraient pas acquises au même

² Certains suffixes permettent la construction de mots nouveaux sans que la catégorie grammaticale de la base ne soit changée (ex. : *plume* – *plumage*). Pour de plus amples détails, se référer à Riegel (1994).

rythme : les règles de flexion seraient apprises beaucoup plus rapidement que celles de la dérivation (Carlisle, 1988).

En ce qui concerne le traitement de la dérivation, sujet de notre recherche, on a remarqué de nombreuses constantes chez les élèves du primaire. Tout d'abord, leur habileté à analyser les morphèmes ainsi que leur expérience de lecture seraient des facteurs déterminants dans la reconnaissance de mots dérivés (Carlisle et Katz, 2006). De plus, les élèves du primaire feraient davantage d'erreurs quand il y a un changement d'orthographe ou un changement phonétique provoqué par l'ajout d'un affixe (Carlisle, 1988) et utiliseraient la stratégie de l'analyse de la racine et des suffixes pour comprendre des mots non familiers en lecture (Nagy et Anderson, 1984; Carlisle, 1988).

Du côté des adultes, toujours en L1, des chercheurs ont observé que tous les participants adultes approchaient les mots morphologiquement complexes de la même manière, c'est-à-dire d'un oeil très analytique. Si le mot morphologiquement complexe est transparent, comme le mot *familial*, ce mot est analysé d'après ses constituants, c'est-à-dire sa racine et ses affixes. Par contre, si le mot morphologiquement complexe est opaque, comme le mot *destruction*, il est plutôt analysé comme un mot simple indécomposable (Dominiek, 1990; Wang, Peng, Guan et Kuang, 1999). Arrivant à une conclusion similaire, Colé et coll. (1989), dans une étude effectuée auprès d'adultes locuteurs natifs du français, ont poussé plus loin cette observation et ils ont constaté que les mots suffixés étaient tous décodés à partir de la racine et qu'ensuite, les membres de la même famille étaient examinés mentalement un par un selon un ordre respectant leur fréquence.

La plupart des observations que nous venons de citer ont été validées auprès de populations anglophones. D'autres chercheurs (Ku et Anderson; Koda; etc.) se sont intéressés au développement morphologique chez des locuteurs parlant le mandarin comme langue première. Dans un premier temps, ils ont eux aussi constaté que les connaissances morphologiques se développaient progressivement

avec l'apprentissage de l'écriture et de la lecture (Koda, 1999). Par contre, dans un deuxième temps, ils ont remarqué que les locuteurs du mandarin acquéraient beaucoup plus tardivement la morphologie dérivationnelle que les locuteurs de l'anglais et ce, en raison de la nature même de leur système d'écriture (Ku et Anderson, 2003). Selon Seidenberg & Gonnerman (2000), les langues varieraient considérablement au point de vue morphologique et ce serait la nature même du système d'écriture de la langue qui en serait la cause, à tout le moins pour la morphologie dérivationnelle (Koda, 1999). En apprenant à écrire et à lire la langue qu'ils parlent, les enfants se familiarisent avec les divers mécanismes de formation des mots. Par exemple, alors que le locuteur d'une langue alphabétique, comme le français et l'anglais, va développer progressivement à l'écrit l'habileté d'analyser la structure interne des mots afin d'en saisir le sens, le locuteur d'une langue logographique, comme le mandarin, va plutôt apprendre à miser sur l'activation du mot entier pour en décoder le sens. Donnons un autre exemple : le locuteur d'une langue qui comporte beaucoup de suffixes à l'écrit, comme le français par exemple, va acquérir, d'une manière consciente ou non, des connaissances morphologiques en lien avec ces suffixes, leur rôle ainsi que leur sens. D'un autre côté, le locuteur d'une langue ayant peu ou pas de suffixes à l'écrit, comme le mandarin par exemple, ne va pas développer les mêmes connaissances morphologiques que notre locuteur francophone. Est-ce que cette différence peut avoir un impact sur la manière dont ces mêmes locuteurs traiteront les mots dérivés dans une L2 ? Nous tenterons de répondre à cette question dans la prochaine section.

1.1.2 La morphologie dérivationnelle en langue seconde

L'acquisition de la morphologie en général et son traitement en L2 est un domaine de recherche très récent (Meara, 1997, pour une revue de la question). Depuis, diverses questions ont été étudiées, notamment concernant l'acquisition de la

morphologie dérivationnelle chez des apprenants adultes. En voici quelques exemples :

- Est-ce qu'une personne bilingue possède un seul lexique mental ou deux lexiques plutôt indépendants ? (Green, 1986; Colomé, 2001)
- Est-ce que la langue maternelle de l'apprenant influence la manière dont les mots de sa L2 seront organisés dans son lexique mental ? (Swan, 1997)
- Est-ce que la L1 joue un rôle dans le traitement des mots complexes de la L2 ? Quelle est l'importance de ce rôle ? (Koda, 1999)
- Est-ce que tous les apprenants adultes et ce, peu importe leur bagage linguistique, ont développé les mêmes connaissances morphologiques ? Par exemple, sont-ils tous capables d'analyser la structure interne des mots, de les décomposer ? (Taft, 2002)
- Est-ce que tous les apprenants décodent les mots dérivés de la même manière? Utilisent-ils tous les mêmes stratégies pour décoder des mots complexes ? Pouvons-nous attribuer certaines stratégies à une famille de langues spécifique? (Vannest et al., 2002)

Dans les prochaines sections, nous discuterons de l'organisation du lexique mental bilingue (voir 1.1.3) ainsi que de l'influence de la langue maternelle sur a) le traitement de la morphologie dérivationnelle (voir 1.1.4) et b) les stratégies d'apprentissage (voir 1.2). Mais tout d'abord, nous souhaitons présenter ici l'état des connaissances que nous avons actuellement sur le traitement de la morphologie dérivationnelle en L2.

Comme nous l'avons mentionné plus tôt, la capacité d'analyser la structure interne de mots complexes est une habileté essentielle lorsque l'on apprend le français, étant donné que plus de sept mots sur dix sont morphologiquement

complexes (Rey-Debove, 1984). Non seulement cette habileté permet à l'apprenant de comprendre des mots complexes, mais elle peut aussi faciliter « the recognition of a new word and its subsequent production » (Laufer, 1997, p. 146). La problématique en lien avec notre recherche est de savoir si tous les apprenants du français langue seconde, peu importe leur bagage linguistique, ont développé cette habileté dans leur L1. Si oui, peuvent-ils la transférer dans leur L2 ? Si non, peuvent-ils l'acquérir dans leur L2 ?

Diverses recherches, dont celle de Koda (1999), concluent qu'il est possible que deux locuteurs n'aient pas développé les mêmes connaissances morphologiques et ce, en raison de la nature du système d'écriture de leur L1 (voir discussion point 1.1.1). Cela signifie, entre autres, que ces deux locuteurs n'approcheront pas la L2 de la même manière. Quelles sont les conséquences d'une telle différence au niveau du traitement des mots de la L2 ?

Selon divers auteurs (Hudelson, 1981; Langer, Bartolome, Vasquez et Lucas, 1990; Nagy, Garcia, Durgunoglu et Hancin-Bhatt, 1993), quand la L1 et la L2 ont plusieurs points en commun au niveau des habiletés et des connaissances générales en lecture, comme les connaissances morphologiques pour décoder le sens de nouveaux mots par exemple, un certain transfert peut s'effectuer de la L1 à la L2. Nous discuterons plus en détails de ce sujet quand nous aborderons la section concernant l'influence de la L1 sur le traitement des mots dérivés de la L2 (point 1.1.4.2).

Par contre, il est également possible que deux langues aient des différences morphologiques très marquées à cause de leur système d'écriture et c'est le cas, entre autres, du mandarin et de l'anglais (Ku et Anderson, 2003). En anglais, il est possible de créer de nouveaux mots à partir de procédés de dérivation et de composition. On retrouvera donc à l'écrit de nombreux mots composés et de mots dérivés. En mandarin par contre, on crée surtout de nouveaux mots en combinant des racines car il y a très peu d'affixes dérivationnels dans cette langue et ceux qui

existent sont peu productifs. L'écrit reflète cette caractéristique de la langue et on y retrouvera surtout des mots composés. De plus, parce que 89% des caractères chinois représentent des morphèmes uniques, ces caractères fournissent des indices visuellement très clairs et très efficaces pour découper les mots polysémiques. L'analyse interne des mots n'est donc pas une pratique enseignée puisque non nécessaire. Finalement, contrairement à l'anglais, les nouveaux mots formés par des procédés de composition impliquent peu de changements d'orthographe ou de phonétique (Ku et Anderson, 2003).

Cette différence entre le mandarin et l'anglais aurait un impact non négligeable sur la performance lors de tâches de reconnaissance de mots. Les résultats de plusieurs recherches (Schmitt et Meara, 1997; Muljani et al., 1998; Koda, 1999; Taft, 2002; Wang, Koda et Perfetti, 2003), que nous présenterons au point 1.1.4.2, ont clairement démontré que cette hypothèse pouvait être confirmée. Selon ces chercheurs, lorsque la L1 et la L2 ont des différences morphologiques très marquées à l'écrit, des difficultés surviennent lors du traitement des mots de la L2, soit parce que le locuteur essaie quand même d'utiliser les connaissances morphologiques qu'il a acquises en L1 et ce, même si elles ne sont pas transférables (Meara, 1984); soit parce que le locuteur n'a tout simplement pas de ressources pour traiter les mots de la L2.

Toutefois, il semblerait que ces difficultés s'amenuiseraient au fur et à mesure qu'augmenteraient les compétences de l'apprenant dans sa L2 (Ellis, 1997). En étant en contact avec le nouveau système d'écriture, l'apprenant se familiariserait avec de nouvelles connaissances morphologiques et pourrait, ainsi, développer de meilleures stratégies pour décoder les mots dérivés de la L2.

Finalement, mentionnons que la L1 n'est pas la seule responsable des erreurs rencontrées lors du traitement de la morphologie en L2. Selon Swan (1997), lorsque deux langues sont dissemblables, beaucoup plus d'erreurs peuvent survenir à cause des difficultés intrinsèques de la L2, réduisant ainsi le rôle de

l'interférence de la L1. Dans le même ordre d'idée, Laufer (1997) a constaté que le manque de régularité dans les combinaisons racine/suffixe était une source de difficulté importante pour les apprenants de la L2 et que cette difficulté ne pouvait pas être attribuée à la L1, mais bien à la nature des règles de la L2.

On connaît encore très peu de choses concernant le traitement des mots complexes en langue seconde. Il en est de même pour le fonctionnement du lexique mental chez les bilingues. Dans la prochaine section, nous présenterons quelques modèles théoriques qui se sont intéressés à la question de l'accès et du fonctionnement du lexique mental.

1.1.3 Modèles théoriques du traitement de la morphologie en L2

Ce n'est que depuis une dizaine d'années que l'on s'interroge sur l'organisation et l'accès au lexique mental des bilingues. Si certaines études ont déjà permis de soulever le voile sur son organisation, comme c'est le cas, par exemple, des études qui démontrent maintenant que les personnes bilingues ne peuvent pas désactiver une de leurs langues lors d'exercices de compréhension (De Bruijn et al., 2001; De Groot et Nas, 1991; De Groot et al., 2000; Dijkstra et al., 2000; Van Heuven et al. 1998; Von Studnitz et Green, 2002), d'autres aspects du lexique mental continuent encore de faire l'objet de débat et c'est le cas, entre autres, du mécanisme même d'activation du lexique mental chez les bilingues.

Dans une tâche de reconnaissance visuelle par exemple, comment les mots sont-ils activés dans le lexique mental bilingue ? Est-ce que tous les mots sont activés en même temps ? Est-ce que certains mots sont plus activés que d'autres ? Qu'est-ce qui détermine le degré d'activation ? Quelques modèles théoriques ont été élaborés afin de tenter d'expliquer ce mécanisme d'activation. En 1987, MacWhinney présente « The Competition Model ». Selon ce modèle, à la vue d'un mot, peu importe la langue, l'apprenant perçoit des signaux et ce sont ces

signaux qui activent des associations. Le signal le plus fort détermine l'association gagnante.

Un autre modèle semblable a été proposé en 2000 par Seidenberg et Gonnerman. Ce modèle, appelé « The Convergence of Codes », explique que la compréhension des mots morphologiquement complexes est attribuable à l'activation simultanée de plusieurs codes (orthographique, phonétique et sémantique) qui vont tous converger vers une association entre une forme et un sens.

En 2002, Dijkstra et Van Heuven proposent à leur tour un modèle théorique pour expliquer le mécanisme d'activation des mots du lexique mental chez les bilingues, « The Bilingual Interactive Activation Plus Model (BIA+) ». Nous avons choisi ce modèle puisqu'il s'est inspiré à la fois des résultats d'études en psycholinguistique, d'études sur le lexique mental des bilingues et des théories proposées en acquisition des langues secondes. En voici sommairement le résumé.

1.1.3.1 The Bilingual Interactive Activation Plus Model (BIA+)

Le modèle BIA+ est d'abord et avant tout une modification du modèle théorique « The Bilingual Interactive Activation Model (BIA) » élaboré en 1998 par Dijkstra et Van Heuven dans le but de fournir une description du système de reconnaissance de mots écrits chez les bilingues. Les modifications au modèle original ont été apportées à la suite de résultats obtenus lors de diverses recherches empiriques, permettant ainsi de mieux comprendre l'architecture du mécanisme d'activation des mots dans le lexique mental des bilingues. Le présent modèle fut présenté dans un article publié par Dijkstra et Van Heuven en 2002 intitulé « *The Architecture of the bilingual word recognition system : From identification to decision.* »

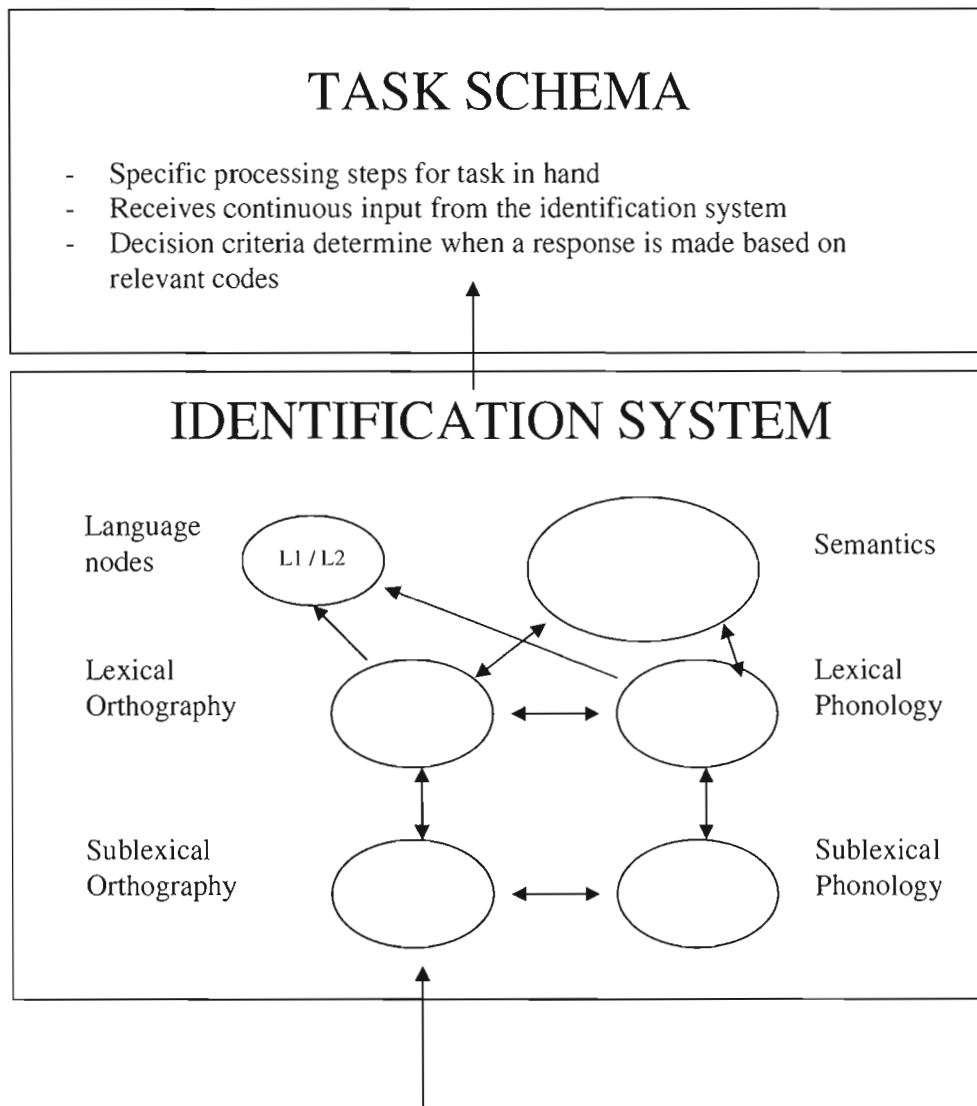


Figure 1.1 The Bilingual Interactive Activation Plus Model. (Tiré de Dijkstra et Van Heuven, 2002). Les flèches indiquent le sens dans lequel l'activation se produit entre les groupes représentés.

Le modèle BIA+ repose sur les principes théoriques suivants :

- Le lexique mental des bilingues est un lexique intégré, unique. Les langues connues n'ont pas leur propre lexique.

- Ce ne sont pas seulement les représentations orthographiques qui permettent une activation des mots, mais plutôt une interaction entre les représentations orthographiques, phonologiques et sémantiques.
- L'activation des mots du lexique mental n'est pas déterminée à l'avance par la langue dans laquelle se déroule l'activité. La sélection de la langue ne se fait que très tardivement, une fois que tous les mots connus et répondant aux critères orthographiques, phonologiques et sémantiques ont été activés.
- Le système de reconnaissance des mots est divisé en deux procédés bien distincts : le système d'identification des mots et le système tâche/décision, c'est-à-dire tout ce qui est en lien avec la tâche, soit les instructions, les attentes du participant, la prise de décision, les éléments et stratégies qui ont permis la prise de décision, etc.
- Le contexte linguistique, c'est-à-dire les sources orthographiques, phonologiques et sémantiques, n'affecte que le système d'identification des mots.
- Le contexte non linguistique, c'est-à-dire les instructions, les objectifs de la tâche, les attentes du participant, n'affecte que le système tâche/décision.
- Ce modèle tente d'expliquer la reconnaissance des mots. Néanmoins, il pourrait également s'appliquer à leur production (p.193).

Avant d'expliquer le fonctionnement d'un tel modèle, il est important de mentionner ici que le BIA+ a été créé dans le but d'expliquer la manière dont les bilingues procèdent pour identifier des mots isolés. L'une des faiblesses de ce modèle, et les auteurs en conviennent, est de ne pas élaborer davantage sur la reconnaissance des mots en contexte. Nous y reviendrons au point 1.1.4.1.

Selon Dijkstra et Van Heuven, ce n'est pas le mot qui est immédiatement perçu par le système d'identification des mots, mais plutôt la série de lettres qui le compose. Par conséquent, dans un premier temps, tous les mots connus par le sujet qui commencent par cette série de lettres spécifique sont activés. Si les langues connues sont très similaires au point de vue orthographique, de nombreux

compétiteurs seront alors activés et tout le mécanisme d'activation sera ralenti. Par contre, si le mot contient une série de lettres spécifique à une langue bien précise, la reconnaissance du mot se fait alors très rapidement.

Une fois les représentations orthographiques activées, elles activent les représentations phonologiques et sémantiques qui y sont associées. Celles-ci vont projeter à leur tour l'information activée vers les « language nodes », qui agissent, en quelque sorte, comme sélecteurs de la langue dans laquelle la réponse doit être formulée. Ensuite, avec toutes les informations recueillies, le système d'identification des mots est prêt à envoyer des réponses au système tâche/décision.

Le système tâche/décision ne reçoit que l'output du système d'identification des mots. Parmi tout l'output reçu, le système tâche/décision doit choisir la bonne réponse. Le choix sera influencé par les instructions et les objectifs de la tâche, les attentes du participant, les stratégies qu'il compte utiliser pour répondre à la question, etc. Selon les auteurs, le participant n'aurait le contrôle que sur le système tâche/décision (p.192). Par exemple, si le système d'identification des mots ne fournit qu'une seule réponse au système tâche/décision, le sujet n'aurait que deux choix : choisir ce mot ou tout simplement ne pas donner de réponse, étant donné qu'il n'a aucun contrôle sur le système d'identification des mots. Autre exemple : le sujet pourrait connaître et utiliser de bonnes stratégies pour décoder des mots complexes, mais ne pas être capable de donner une réponse si son système d'identification des mots ne peut lui en fournir.

Bien que les concepteurs de ce modèle aient tenté d'inclure le plus de données connues possibles, de nombreuses interrogations planent toujours. Par exemple, certains auteurs se demandent si le rôle du décodage phonétique n'a pas été minimisé dans le modèle (Brysbaert, Van Wijnendaele et Duyck, 2002). D'autres aimeraient savoir où situer les différences individuelles (Van Hell, 2002). Dans le cadre de ce mémoire, nous avons cherché à connaître la manière dont ce modèle

explique l'influence de la L1 au niveau du décodage des mots de la L2, si influence il y a. Ce sujet sera l'objet de la prochaine section.

1.1.4 Influence de la L1 dans le traitement des mots dérivés de la L2

Dans la présente section, nous aborderons d'abord la manière dont le modèle BIA+ explique l'influence de la L1 dans le traitement des mots de la L2 et dans un deuxième temps, nous présenterons les résultats d'études qui se sont intéressées à cette question.

1.1.4.1 Le modèle BIA+ et le rôle de la langue première

Selon le modèle BIA+, lorsqu'un apprenant procède à une tâche de reconnaissance de mots, tous les mots de son lexique mental correspondant aux signaux envoyés sont activés et ce, peu importe la langue. La sélection de la langue ne se fait qu'à la toute fin dans le processus d'activation. Plus il y a de ressemblances entre les mots des deux langues, d'un point de vue orthographique surtout, plus il y a de compétiteurs activés. À l'inverse, moins il y a de ressemblances entre les deux langues, moins il y a de compétiteurs activés (Dijkstra et Van Heuven, p.183). Dans cette perspective, la L1 joue un rôle au niveau du temps de réponse, le nombre de congénères influençant directement la vitesse à laquelle le système d'identification des mots traitera l'information.

Comme nous l'avons mentionné plus tôt, les mots des deux langues correspondant aux signaux sont activés simultanément. Toutefois, comme le mentionnent les auteurs du modèle BIA+, dans le cas où un mot isolé présenterait des caractéristiques propres à l'une des deux langues, « we will (of course) find evidence of language specific access (e.g., Chinese orthography will not induce much Latin letter activation) » (p.183). Encore une fois, la L1 n'influencerait que le temps de réponse.

Le modèle BIA+ ne s'adresse qu'aux apprenants bilingues en bonne possession de deux ou de plusieurs langues, mais qu'en est-il des apprenants débutants, intermédiaires et avancés d'une langue seconde dont les connaissances linguistiques ne sont pas comparables à celles d'un locuteur natif ? Est-ce que la L1 aurait un rôle plus important dans le modèle si celui-ci avait été pensé pour eux ? Selon Van Hell (2002), beaucoup d'études suggèrent que les différences individuelles des apprenants, et même les différences individuelles des bilingues, non seulement au niveau de la compétence, mais également au niveau du rythme d'apprentissage et de l'expérience de lecture, « may modulate cross-language processing » (p.211). Malheureusement, les études sur le sujet ne permettent pas encore de bien saisir ces variations ni d'identifier s'il y a une réelle influence de la L1 dans ces contextes.

Finalement, puisque le modèle BIA+ a été conçu, d'abord et avant tout, pour expliquer les mécanismes d'activation du lexique mental lors de tâches de reconnaissance de mots écrits isolés, on sait encore très peu de choses quand la reconnaissance de mots se fait en contexte. Sans trop s'aventurer, les auteurs du modèle mentionnent que le système d'identification des mots, en plus d'envoyer de l'output au système tâche/décision, interagit avec le système « analyse de phrase ». De quelle manière interagissent-ils ? Interagissent-ils avant que le système d'identification des mots envoie des réponses au système tâche/décision ? Les auteurs donnent très peu de détails. Selon eux, alors que la reconnaissance de mots écrits isolés active d'abord les représentations orthographiques, la reconnaissance de mots en contexte alimenterait plutôt les représentations syntaxiques et sémantiques dans les deux langues (propos rapportés dans Van Hell, 2002). Malgré ces explications, une question demeure : serait-il possible qu'il y ait une préactivation des « language nodes » lorsque le mot à reconnaître s'inscrit dans un contexte langagier précis ? Très peu d'études jusqu'à présent se sont intéressées à ce sujet (Dijkstra & Van Heuven, 2002). Toutefois, celles qui s'y sont intéressées en ont tiré deux conclusions :

a) « Lexical form features are more important than language membership in the determination of word recognition in sentences » (Altarriba, Kroll, Sholl & Rayner, 1996).

b) « (...) orthographic representations, in all activations, play a major role » (Pexman et Lupker, 2001)

Ainsi, l'activation de mots de la L2 qui auraient des caractéristiques particulières, comme beaucoup d'infexes par exemple, une caractéristique que l'on ne retrouverait pas dans les mots de la L1, réduirait considérablement le nombre de compétiteurs. De plus, selon la deuxième conclusion, les représentations orthographiques joueraient un rôle de premier plan dans le mécanisme d'activation des mots alors que les indices sémantiques ou syntaxiques fournis par le contexte seraient plutôt secondaires. Cette dernière conclusion est similaire à celle proposée par Schwartz et Kroll (2006) dans une étude menée auprès de participants bilingues parlant l'espagnol et l'anglais. Selon eux, on ne peut pas vraiment affirmer que le contexte fournit un indice assez important sur la nature de la langue pour qu'il y ait une préactivation des « language nodes » (p.18).

En résumé, dans le modèle BIA+, la L1 joue surtout un rôle au niveau du temps de réponse. Plus il y a de compétiteurs lorsque les deux langues connues ont des similitudes au niveau de l'orthographe, par exemple, plus le temps de réponse sera long. Par contre, moins il y a de compétiteurs lorsque les langues connues ont peu de similitudes au niveau de l'orthographe, et plus court sera le temps de réponse. Il est important, toutefois, d'indiquer ici que le temps de réponse et la qualité des réponses sont deux choses bien distinctes. Le temps de réponse peut être très long et la réponse donnée mauvaise alors que le temps de réponse peut être très court et la réponse donnée également erronée.

Bien que ce modèle ne soit pas parfait, il ne faut surtout pas oublier qu'il s'agit d'un modèle théorique et donc, sujet à de nombreux changements.

Passons maintenant à la présentation de diverses recherches qui ont tenté de vérifier s'il y a effectivement une influence non négligeable de la L1 au niveau du traitement de la morphologie en L2.

1.1.4.2 Recherches sur l'influence de la L1 dans le traitement de la morphologie en L2

Depuis quelques années, des chercheurs ont observé que la L1 d'un apprenant avait une certaine influence sur le décodage des mots de la L2 (Frost, Katz et Bentin, 1987; Hatta, 1992; Henderson, 1984, 1985; Nakagawa, 1994; Rusted, 1988; Turvey, Feldman et Lukatela, 1984; Tzeng et Wang, 1983; Ellis, 1997; Koda, 1999). On a tenté, par la suite, de préciser cette observation en spécifiant les cas où cette influence se produisait. Les résultats que nous présenterons maintenant indiquent que les différences au niveau du décodage des mots complexes en anglais langue seconde³ se produisent surtout lorsque la L1 du sujet est distante de la L2 du point de vue morphologique.

Plusieurs chercheurs s'entendent pour dire que la L1 a une influence considérable sur la manière dont la L2 sera apprise et utilisée (Kellerman, 1984; Kellerman et Smith, 1986; Ringbom, 1987; Odlin, 1989; Perdue, 1993). Quel type d'influence a-t-elle ? Afin de répondre à cette question, Taft (2002) a mené une expérience auprès de Japonais et de Chinois apprenant l'anglais. Il voulait savoir comment des mots anglais polysyllabiques étaient analysés par de tels apprenants. Ses résultats démontrent que les Japonais et les Chinois n'analysent pas les mots polysyllabiques anglais de la même manière, les Chinois éprouvant davantage de difficultés. Selon Taft, l'analyse des mots polysyllabiques en anglais L2 est

affectée par la manière dont ces mots sont structurés à l'écrit dans la L1. Les Chinois, comme nous l'avons mentionné précédemment, utilisent des logogrammes, c'est-à-dire des caractères qui représentent, dans la majorité des cas, des mots complets. Les Japonais, quant à eux, ont deux ensembles de caractères distincts : les *kanji* et les *kana*. Alors que les *kanji* sont des logogrammes d'origine chinoise qui ont été modifiés, les *kana* sont plutôt des syllabaires. On les appelle des syllabaires puisque ce sont des syllabes qui sont représentés par des caractères et non des mots entiers, et c'est la combinaison de ces syllabes qui forment les mots. Les *kana* permettent de noter quasiment phonétiquement des mots, des affixes grammaticaux ainsi que des mots d'origine étrangère. Cette notion de combinaison de syllabes n'étant pas connue (ou maîtrisée) par les Chinois, contrairement aux Japonais, il est possible qu'ils aient tout simplement éprouvé plus de difficultés à procéder à l'analyse de mots polysyllabiques. Selon Taft, plus la L1 est distante de la L2 du point de vue morphologique, plus l'analyse des mots de la L2 sera difficile.

Même observation du côté de Schmitt et de Meara (1997) qui ont entrepris une étude auprès de sujets adultes japonais apprenant l'anglais. Ils ont constaté que le peu de connaissances des suffixes qu'ont les Japonais venait influencer leur capacité à traiter la morphologie en anglais L2. Cette constatation vient, encore une fois, corroborer l'hypothèse que plus la L1 est différente de la L2 du point de vue morphologique, et plus il est difficile de décoder des mots complexes dans la L2.

Koda (1999), dans une expérience menée auprès d'adultes coréens et chinois apprenant l'anglais langue seconde, a également constaté une influence de la L1 au niveau de la sensibilité à la structure interne des mots complexes de la L2. Selon Koda, les connaissances morphologiques acquises dans la L1, comme la

³ Jusqu'à présent, très peu d'études se sont intéressées au traitement de la morphologie dérivationnelle en français langue seconde.

sensibilité à la structure interne des mots, sont directement transférées dans la L2⁴. Les difficultés surviennent lorsque la L1 et la L2 ne requièrent pas les mêmes connaissances morphologiques pour le décodage. Koda en est arrivé à cette conclusion en constatant que les Chinois éprouvaient davantage de difficultés à décoder les mots anglais que les Coréens, qui ont un système d'écriture dont la morphologie est semblable à celle de la langue anglaise. Toutefois, Koda précise que si la L1 provoque des variations au niveau de la performance lors du décodage de mots anglais, ces variations ne se traduisent pas nécessairement, du côté quantitatif, en une moins bonne performance.

En résumé, selon Koda, le système d'écriture de la L1 et les connaissances morphologiques que ce système permet de développer serait le grand responsable des différences observées au niveau du traitement des mots dérivés de la L2⁵. Cette conclusion a été appuyée par les résultats de recherches similaires. Par exemple, Wang, Koda et Perfetti (2003), dans le cadre d'une étude presque identique, en sont arrivés à la conclusion que : « Cross-writing system differences in L1s and L1 reading skills transfer could be responsible for ESL performance differences in English word identification task » (p. 129). Wang et Koda (2005) ont pu observer qu'une L1 alphabétique facilitait la reconnaissance de mots d'une L2 également alphabétique, alors qu'une L1 non alphabétique ne facilitait pas la reconnaissance de mots d'une L2 alphabétique. Même observation du côté de Muljani et al. (1998) qui ont comparé les résultats d'apprenants indonésiens et chinois lors de tests de reconnaissance de mots anglais. Ils ont constaté que plus la L1 et la L2 ont des systèmes d'écriture différents, et plus le traitement des mots de la L2 est difficile. Lors de leur recherche, ils ont remarqué que les apprenants indonésiens, dont la L1 est alphabétique, avaient mieux performé en général lors de tests de reconnaissance de mots anglais, langue également alphabétique, que les Chinois, dont la L1 est non alphabétique.

⁴ Ces connaissances, toutefois, ne sont pas une garantie de succès car encore faut-il que le sujet ait les habiletés pour utiliser ces connaissances lors du traitement des mots.

⁵ Pour de plus amples explications, se référer au point 1.1.1.

Mentionnons, pour terminer, que si les connaissances morphologiques de la L1 ont un impact non négligeable sur le traitement des mots de la L2, il semblerait que les stratégies d'apprentissage apprises dans la L1 auraient également une influence dans le traitement des mots de la L2 (Gass & Selinker, 2001). Nous aborderons le thème des stratégies d'apprentissage dans la prochaine section.

1.2 Les stratégies d'apprentissage

Depuis plus de trente-cinq ans, des chercheurs s'intéressent à la question des stratégies d'apprentissage (voir Schmitt, 1997, pour une revue de la question). Qu'est-ce qu'une stratégie d'apprentissage ? Dans le cadre de notre recherche, la définition qui nous semble la plus appropriée est celle donnée par Faerch et Kasper (1980) : une stratégie est « (...) a potentially conscious plan for solving what, to the individual, presents itself as a problem in reaching a particular goal » (p.60). À cette définition, nous aimerions ajouter que, tout comme Feldmann et Stemmer (1987), nous croyons qu'une stratégie n'est pas figée, mais bien dynamique et qu'elle peut interagir avec une autre stratégie lors d'exercices.

Depuis que l'on s'intéresse aux stratégies d'apprentissage, plusieurs classifications ont été proposées (Bialystok, 1981; O'Malley et Chamot, 1990; Oxford, 1990; etc.). Dans la prochaine section, nous présenterons d'abord quelques-unes de ces classifications (1.2.1) avant de nous pencher sur les stratégies d'apprentissage en lien avec la morphologie (1.2.2) et les recherches qui se sont intéressées à l'influence de la L1 sur celles-ci (1.2.3).

1.2.1 Taxinomies

Choisir une classification des stratégies d'apprentissage n'est pas chose facile car de nombreuses listes ont été compilées et ce, basées sur des objectifs de recherche

bien divers. Le deuxième but de notre recherche étant d'identifier les stratégies utilisées lors d'un exercice de décodage et de production de mots morphologiquement complexes, il nous fallait trouver une classification qui :

- a) accorderait une place importante aux stratégies en lien avec la morphologie
- b) présenterait des stratégies reliées à la L1.

Nous avons consulté plusieurs taxinomies dont, entre autres, la classification proposée par Schmitt (1997) et celle de Feldmann et Stemmer (1987). La taxinomie de Schmitt est une liste exhaustive de toutes les stratégies identifiées jusqu'à présent (lire ici jusqu'en 1997). Cette liste, bien que complète, est beaucoup trop générale pour convenir aux besoins précis de notre recherche.

La classification des stratégies d'apprentissage présentée par Feldmann et Stemmer dans un article intitulé « *Thin_ aloud a_ retrospective da_ in C-te_ taking : diffe_ languages _ diff_ learners_ s_ approaches?* » (1987) est intéressante puisqu'elle présente, entre autres, des stratégies qui ont été identifiées par des apprenants lors d'un exercice de décodage mettant en jeu les connaissances morphologiques. Bien que cette classification se rapproche davantage des objectifs de notre recherche, elle présente toutefois un inconvénient majeur : elle n'identifie pas clairement une stratégie en lien avec la morphologie dérivationnelle. L'auteur mentionne bien la possibilité que le sujet essaie de traduire les préfixes, mais cette stratégie fait partie d'une catégorie beaucoup plus vaste appelée « recall by search of meaning - translation to mother tongue » (p.260).

Dans le cadre de cette recherche, nous nous inspirerons de la classification élaborée par Haastруп (1987) à partir de données qu'elle a obtenues auprès de 62 étudiants danois apprenant l'anglais langue seconde. L'originalité de sa classification réside dans le fait que ce ne sont pas des stratégies qui ont été répertoriées, mais bien les indices que les élèves ont utilisés pour décoder le sens des mots inconnus.

Selon Haastrup, les apprenants qui doivent deviner le sens de mots nouveaux dans leur L2 font appel à trois types d'indices :

- « Interlingual : cues based on L1, loanwords in L1 or knowledge of foreign languages other than English
- Intralingual : cues based on knowledge of English
- Contextual : cues based on the text or on informants' knowledge of the world » (p.199)

Ce sont un ou plusieurs de ces indices qui aident l'apprenant à décoder les mots inconnus de la L2. Regardons plus en détails ces indices.

Tableau 1.1
Taxinomie des indices servant à décoder des mots nouveaux dans la L2.
(Tiré de Haastrup, 1987, p. 199.)

Taxonomy of knowledge sources		
Contextual	Intralingual	Interlingual
a) The text <ul style="list-style-type: none"> • A single word from the immediate context • The immediate context • A specific part of the context beyond the sentence of the test word • Global use of the text 	a) The test word <ul style="list-style-type: none"> • Phonology / orthography • Morphology • Prefix • Suffix • Stem • Lexis • Word class • Collocations • Semantics 	a) L1 <ul style="list-style-type: none"> • Phonology / orthography • Morphology • Lexis • Collocations • Semantics
b) Knowledge of the world	b) The syntax of the sentence	b) Lⁿ (Latin, German, French, etc.) <ul style="list-style-type: none"> • General reflections • Reflections about the origin of the word • Test word pronounced in Lⁿ • Morphology • Lexis • Semantics

Cette liste présente non seulement l'avantage de répondre à nos objectifs de recherche, mais elle laisse également une place à la possibilité que l'apprenant se soit fié à des indices provenant d'une langue autre que sa langue maternelle pour décoder des mots. Cet aspect ne doit pas être négligé lorsque l'on travaille auprès d'adultes provenant de différents horizons linguistiques.

Finalement, nous nous référerons également à une autre taxinomie lorsqu'il sera question d'interpréter nos résultats. Complémentaire à la liste proposée par Haastруп, la classification de Gu et Jonhson (1996) puise son originalité dans le fait qu'il s'agit d'une liste de stratégies que 850 apprenants chinois ont eux-mêmes classifiées, non pas suite à un exercice, mais plutôt selon une auto-évaluation. Dans un premier temps, ces étudiants ont évalué l'utilité de chaque stratégie qui leur a été présentée selon une échelle de 1 à 7. Dans un deuxième temps, les deux chercheurs ont calculé la moyenne (M) des notes obtenues pour chaque stratégie. Cela leur a permis d'obtenir une idée globale des préférences de l'ensemble des étudiants chinois. Voici les résultats obtenus par Gu et Johnson:

Tableau 1.2
Learning strategies : self-reports.
(Tiré de Gu et Johnson, 1996, p. 653.)

How Chinese Learners Learn Vocabulary : Self reports			
Categories and Strategies	M	SD	N
Beliefs			
• Words should be memorised	3.04	0.83	849
• Acquire vocabulary in context	4.94	0.78	850
• Learn vocabulary and put it to use	5.74	0.62	847
Metacognitive Regulation			
• Selective attention	4.23	0.86	822
• Self-initiation	4.58	1.00	829
Guessing Strategies			
• Wider context	4.60	0.85	824
• Immediate context	4.47	0.84	824
Dictionary Strategies			
• Comprehension	4.97	1.00	830
• Extended dictionary strategies	4.82	0.93	820
• Looking-up strategies	4.55	0.94	823

Note-Taking Strategies			
• Meaning-oriented note-taking	4.15	0.99	824
• Usage-oriented note-taking	4.27	1.14	824
Rehearsal Strategies			
• Using word lists	3.15	0.99	824
• Oral repetition	4.20	1.07	840
• Visual repetition	3.92	1.17	833
Encoding Strategies			
• Association / elaboration	3.69	0.97	826
• Imagery	3.11	1.00	833
• Visual encoding	4.00	1.08	831
• Auditory encoding	3.69	1.19	832
• Using word-Structure	3.96	1.16	840
• Semantic encoding	3.24	1.03	839
• Contextual encoding	4.11	1.07	829
Activation Strategies	3.80	1.05	827

Dans ce tableau, il est intéressant de constater que la stratégie de l'analyse interne des mots (using word-structure) se classe au troisième rang dans la catégorie « Encoding strategies », tout de suite après les indices contextuels et les indices visuels. Rappelons que ce classement se base sur la perception des étudiants et qu'il ne sera intéressant que dans la perspective où nous pourrons le comparer avec les stratégies que nos participants auront identifiées.

Passons maintenant à la description des stratégies en lien avec la morphologie dérivationnelle.

1.2.2 Les stratégies d'apprentissage et la morphologie dérivationnelle

Dans toutes les classifications (ou presque) mentionnées précédemment, on retrouve la stratégie de l'analyse interne du mot, une stratégie en lien direct avec la morphologie. Cette stratégie, selon la taxinomie, se présente sous diverses appellations : « using word structure » (Gu et Johnson, 1996); « intralingual cues – Morphology : prefix, suffix, stem » (Haastrup, 1987); « analyse affixes and roots » (Shmitt, 1997); « using word parts » (Nation, 1990).

Cette stratégie, peu importe son nom, consiste essentiellement à décomposer le mot en constituants afin de mieux en saisir le sens. Comme le mentionne Nation dans son livre « *Teaching and learning vocabulary* » (1990), afin d'être en mesure d'utiliser cette stratégie, l'apprenant a besoin de trois habiletés très précises :

- 1- L'apprenant doit d'abord être capable de diviser le mot en constituants afin d'en identifier la racine et les affixes.
- 2- L'apprenant doit connaître le sens de chacun des constituants.
- 3- L'apprenant doit être capable d'établir une connexion entre le sens dégagé par les constituants et le sens trouvé dans le dictionnaire.
(p.167)

Sans ces trois habiletés, l'apprenant ne peut pas utiliser efficacement la stratégie de l'analyse des constituants pour reconnaître un mot et en découvrir le sens.

Si ces habiletés se développent naturellement en L1 lorsque, bien sûr, la langue écrite en question contient de nombreux affixes, ce n'est pas le cas lorsque la L1 écrite est plutôt pauvre en affixes. Une personne dont la L1 a peu d'affixes, par exemple, n'aura pas nécessairement développé les trois habiletés requises pour analyser la structure interne des mots, étant donné que cette stratégie est non pertinente dans sa langue. Mais qu'en est-il lorsqu'il est question de l'apprentissage d'une langue seconde ? Est-ce que les stratégies sont transférables d'une langue à une autre ? Est-ce que le fait de ne pas avoir acquis la stratégie de l'analyse interne du mot dans sa L1 nuit au décodage des mots d'une L2 riche en affixes ? Voici les résultats de quelques recherches qui se sont intéressées au sujet.

1.2.3 Recherches sur l'influence de la L1 sur les stratégies d'apprentissage

Les stratégies utilisées pour décoder et produire des mots dans la L2 peuvent être influencées par les stratégies utilisées dans la L1. Heilenman et McDonald (1993), dans une étude menée auprès de sujets adultes apprenant le français,

concluent qu'il y a une nette indication de tentatives de transfert de stratégies de la L1 à la L2. Dans certains cas, les stratégies de la L1 fonctionnent si bien que les apprenants ne cherchent pas à développer de meilleures stratégies beaucoup plus adaptées à la L2. Par contre, dans d'autres cas, les stratégies de la L1 ne sont tout simplement pas transférables, ce qui serait surtout le cas lorsque la L1 et la L2 sont très distantes du point de vue orthographique. Cela occasionnerait de nombreuses difficultés pour l'apprenant qui, malgré tout, « (...) continue to use these strategies for handling words in their L2, even if the strategies are not particularly well adapted » (Meara, 1984, p.234).

Wang et Koda (2005) parviennent à la même conclusion lors d'une enquête effectuée auprès d'étudiants coréens et chinois apprenant l'anglais langue seconde. Selon eux, les Coréens ont obtenu de meilleurs résultats lors de tâches de reconnaissance de mots anglais puisqu'ils ont pu transférer les stratégies de lecture qu'ils avaient développées dans leur L1, contrairement aux apprenants chinois qui n'avaient pas développé de telles stratégies dans leur L1.

Heureusement, il semblerait que plus le niveau de compétence en L2 augmente, et plus les stratégies utilisées par les apprenants ressemblent à celles qui sont propres à la L2 et ce, peu importe la L1 (ce qui diminuerait les difficultés de décodage lorsque la L1 et la L2 sont très distantes morphologiquement) (Miao, 1981; Sasaki, 1991; Hoover et Dwivadi, 1998; Ellis, 1997). Le niveau de compétence dans la L2 ne serait toutefois pas le seul facteur influençant les stratégies. De nombreux autres aspects seraient en jeu lorsqu'il est question de l'acquisition, du transfert et de l'utilisation des stratégies, entre autres, la maturité cognitive (Miramontes et Commins, 1989), le contexte d'apprentissage (O'Malley & Chamot, 1990) et la culture de l'apprenant (Schmitt, 1997)).

Selon Miramontes et Commins (1989), pour qu'il y ait un transfert efficace des stratégies de la L1 vers la L2, il faut qu'il y ait avant tout une certaine maturité

cognitive. Avec cette maturité, l'apprenant est en mesure de juger les stratégies qu'il peut transférer de celles qu'il ne peut pas.

Le contexte d'apprentissage jouerait également un rôle dans l'acquisition des stratégies. Lors d'une recherche portant sur l'enseignement des stratégies reliées au vocabulaire en L2, O'Malley et Chamot (1990) ont pu constater que chez certains apprenants, l'apprentissage systématique de stratégies avait un effet bénéfique sur le transfert. L'enseignement systématique permettrait à l'apprenant de connaître et de s'approprier de nouvelles stratégies plus adaptées à la L2 et de délaisser celles acquises en L1.

Lors de cette même recherche, O'Malley et Chamot ont fait une étonnante découverte concernant l'influence de la culture sur l'acquisition et le transfert de stratégies d'apprentissage : les asiatiques, comparativement aux hispanophones, ont résisté à l'enseignement systématique des stratégies. Cette influence de la culture sur les stratégies a également été observée par Schmitt (1997), dans une recherche similaire. Celui-ci a constaté que les apprenants de différentes cultures pouvaient avoir quelques fois des opinions très différentes sur l'utilité et l'efficacité de certaines stratégies pour apprendre le vocabulaire.

La résistance à l'enseignement systématique ainsi que la diversité des opinions quant à l'utilité des stratégies pourraient, en partie, être expliquées par toute la question de la perception que les apprenants ont de la distance entre leur L1 et leur L2. Comme l'expliquent Ridley et Singleton (1995):

« There is a fair amount of research evidence to suggest that the degree of similarity or divergence that learners perceive between their L1 and their target language will strongly influence the general extent to which they draw on their L1 resources in coming to grips with L2 problems (see, e.g., Kellerman, 1977, 1979; Ringbom, 1987: *passim*; Odlin, 1989:140ff.) » (p.140).

Si l'apprenant croit que sa L1 est très distante de sa L2, peu importe si cette perception est fausse ou réelle, il y a peu de chance pour qu'il y ait des transferts efficaces de stratégies.

En résumé, les recherches démontrent qu'il existe bel et bien un transfert des stratégies de la L1 à la L2 et que ce transfert n'est pas toujours souhaitable, surtout lorsque la L1 et la L2 sont distantes du point de vue morphologique par exemple. Ce problème de transfert semble toutefois se résoudre de lui-même avec la progression de la compétence de l'apprenant dans sa L2 et l'acquisition d'une plus grande maturité cognitive. Néanmoins, certains facteurs comme la culture de l'apprenant et sa perception de la distance entre sa L1 et sa L2 pourraient faire en sorte que l'apprenant résiste à l'apprentissage de nouvelles stratégies et continue d'utiliser les stratégies qu'il a développées dans sa L1 pour résoudre les problèmes de sa L2.

Le cadre théorique et la problématique de notre recherche ayant maintenant été exposés, poursuivons avec la présentation de notre méthodologie.

CHAPITRE II

MÉTHODOLOGIE

2.1 Présentation de la recherche

Les résultats des recherches que nous avons présentés dans le chapitre précédent nous permettent de conclure que les caractéristiques morphologiques du système d'écriture de la langue maternelle d'un apprenant a un impact non négligeable sur le traitement des mots dérivés en anglais langue seconde. Plus le système d'écriture de la langue maternelle de l'apprenant est distant du point de vue morphologique du système d'écriture de l'anglais, plus l'analyse des mots de cette deuxième langue sera difficile pour cet apprenant.

Ces mêmes recherches nous ont également appris que les étudiants avaient tendance à utiliser les stratégies qu'ils avaient développées dans leur langue première (L1) pour solutionner les problèmes rencontrés dans leur langue seconde (L2). Les difficultés survenaient lorsque la L1 était différente du point de vue orthographique de la L2. Toutefois, lorsque le niveau de compétence en langue seconde augmentait, les stratégies utilisées par les apprenants ressemblaient de moins en moins aux stratégies de la L1 et de plus en plus à celles propres de la L2.

Par le biais de cette recherche, nous avons voulu vérifier s'il existait un impact semblable de la langue maternelle dans le traitement des mots dérivés en français langue seconde ainsi qu'au niveau des stratégies employées pour traiter ces mots dérivés. On peut émettre l'hypothèse que les résultats obtenus seront similaires aux résultats obtenus pour l'anglais étant donné que le français est comparable à l'anglais du point de vue de la morphologie. Si cette hypothèse était validée, il serait alors intéressant de la partager avec les enseignants afin qu'ils puissent adapter leur enseignement du vocabulaire aux besoins de leurs apprenants.

Dans le présent chapitre, nous présenterons d'abord nos questions et hypothèses de recherche. Par la suite, nous décrirons dans un premier temps nos participants, dans un deuxième temps nos instruments de mesure et nous expliquerons ensuite le déroulement et la collecte de nos données. Nous terminerons en expliquant la manière dont nous avons procédé pour coder les résultats et préciserons, brièvement, les différentes analyses que nous avons effectuées.

2.1.1 Questions de recherche

Dans le cadre de cette étude, nous avons tenté de répondre aux questions suivantes :

- a) Est-ce que des apprenants adultes d'horizons linguistiques différents réussissent tous avec la même facilité une tâche demandant décodage et production de mots suffixés en français langue seconde ?
- b) Est-ce que des apprenants adultes d'horizons linguistiques différents utilisent les mêmes stratégies pour décoder et produire des mots suffixés en français langue seconde ?

Afin de rendre nos questions de recherche opérationnelles, il a fallu préciser ce que nous entendions par « *adultes d'horizons linguistiques différents* ». Si nous voulions vérifier l'impact de la variable « langue maternelle de l'apprenant » sur a) la performance dans le traitement des mots dérivés et sur b) les stratégies d'apprentissage utilisées lors du traitement de ces mots, il fallait être en mesure de comparer les résultats d'étudiants provenant d'horizons linguistiques différents. À la lumière des études que nous avons présentées au chapitre I, il nous a semblé préférable de regrouper les étudiants selon la richesse morphologique de leur L1 à l'écrit. Nous avons donc créé deux groupes : le premier groupe comprend tous les participants ayant comme langue écrite une langue de type isolant, c'est-à-dire une langue écrite dont les mots ont tendance à être invariables et qui, par

conséquent, comprend peu d'affixations (LPA) ; le deuxième groupe, quant à lui, regroupe tous les participants ayant comme langue maternelle écrite une langue de type flexionnel, soit une langue dont les mots écrits peuvent changer de forme, entre autres, par le procédé d'affixation. Ce type de langue fait appel régulièrement au procédé d'affixation et comprend donc beaucoup d'affixes (LBA). Par rapport à notre échantillon, nous avons regroupé dans la catégorie LPA les étudiants parlant le mandarin. Cette langue est dite « isolante » puisque que la majorité de ses mots sont ou ont tendance à être invariables. De plus, les morphèmes de cette langue sont surtout des morphèmes libres, et non liés, et ils sont la plupart du temps décodés comme des mots graphiques isolables, et non comme des particules qui ont été combinées pour former des mots. Le mandarin contient bien sûr quelques morphèmes liés, dont, entre autres, des suffixes, mais ceux-ci sont beaucoup moins évidents à l'écrit qu'à l'oral (Tiee, 1986). Finalement, ajoutons qu'en raison de son système d'écriture, le mandarin ne permet pas une analyse interne des mots n'ayant qu'un seul morphème (Taft, 2002) ; que c'est une langue qui comporte beaucoup plus de mots composés que de mots dérivés ; et que du côté des suffixes, on retrouve très peu de suffixes productifs (Ku et Anderson, 2003).

Dans la catégorie LBA, nous avons regroupé tous les étudiants de notre échantillon parlant une langue indo-européenne (voir 2.3 pour une description plus détaillée du profil linguistique de nos participants). Toutes les langues indo-européennes sont, à divers degrés, de type flexionnel. Nous ne précisons toutefois pas dans ce mémoire les divers degrés de flexion de chacune de ces langues, le nombre de sujets pour chacune d'entre elles étant trop petit pour que ces détails soient significatifs.

2.1.2 Hypothèses de recherche

Suite à ces précisions, nous avons formulé les hypothèses suivantes :

Étant donné que le français écrit est riche en affixes :

- 1- La performance au niveau du décodage et de la production de mots suffixés en français langue seconde sera moins bonne chez les apprenants dont la langue maternelle écrite a peu d'affixes que chez les apprenants dont la langue maternelle écrite comprend davantage d'affixes.
- 2- Les stratégies employées lors du décodage et de la production de mots suffixés en français langue seconde par des apprenants dont la langue maternelle écrite contient peu d'affixes seront différentes de celles qui sont employées par des apprenants dont la langue maternelle écrite comprend davantage d'affixes.

Afin de confirmer ou d'infirmer ces hypothèses, il nous a fallu choisir des instruments de mesure appropriés. Avant de présenter ceux-ci, passons à la description de nos participants.

2.2 Participants

Les sujets de cette étude étaient des adultes et de jeunes adultes apprenant le français langue seconde dans le cadre du « *Certificat en français écrit pour non-francophones* » offert par l'École de langues d'une université francophone du Québec. Dans la majorité des cas, ces étudiants étaient tous inscrits au certificat à temps plein, soit 12 crédits universitaires à raison de 12 heures de cours par semaine. Nos participants étaient tous des immigrants hautement scolarisés dans leur langue première, qui étaient arrivés au Québec depuis peu (deux mois à un an) et qui souhaitaient y demeurer.

Nous avons évalué trois groupes d'étudiants faisant partis du niveau avancé pour une population totale d'une centaine d'étudiants. Étant administrés lors des heures de classe, les C-tests complétés par les étudiants absents lors de l'une ou de l'autre des épreuves ont dû être rejetés. De plus, les C-tests qui n'ont pas été suffisamment complétés ont également dû être écartés de l'étude. Par conséquent, de la centaine d'étudiants que nous avions au départ, seulement 63 d'entre eux ont complété en bonne et due forme les C-tests. Voici deux tableaux décrivant leur profil. Il est à noter que dans ces tableaux, les étudiants ont été répartis selon la richesse morphologique de leur langue maternelle écrite, à savoir les étudiants qui ont une L1 écrite ayant peu d'affixes (LPA) et ceux ayant une L1 écrite comprenant beaucoup d'affixes (LBA).

Tableau 2.1
Profil des participants / Sexe

	C-tests		Tâche de rappel stimulé	
	<i>Homme</i>	<i>Femme</i>	<i>Homme</i>	<i>Femme</i>
LPA	6	30	2	10
LBA	8	19	3	9
Total	<i>n</i> = 63		<i>n</i> = 24	

Tableau 2.2
Profil des participants / L1

C-tests (<i>n</i> = 63)	Tâche de rappel (<i>n</i> = 24)
LPA	LPA
Mandarin (<i>n</i> = 36)	Mandarin (<i>n</i> = 12)
LBA	LBA
Anglais (<i>n</i> = 1)	Anglais (<i>n</i> = 1)
Bulgare (<i>n</i> = 2)	Bulgare (<i>n</i> = 1)
Espagnol (<i>n</i> = 14)	Espagnol (<i>n</i> = 5)
Roumain (<i>n</i> = 4)	Roumain (<i>n</i> = 1)
Russe (<i>n</i> = 6)	Russe (<i>n</i> = 4)

Avant de passer à la présentation de nos instruments de mesure, nous devons préciser que les vingt-quatre participants qui ont bien voulu se soumettre à la tâche de rappel stimulé font partis des soixante-trois étudiants qui ont complété les deux premiers C-tests. Une description plus détaillée des participants sera présentée lorsque nous aborderons les résultats de notre recherche (voir chap. III).

2.3 Instruments de mesure

Puisqu'il fallait observer à la fois la performance des apprenants dans une tâche de décodage et de production de mots suffixés ainsi que les stratégies employées par ces mêmes apprenants pour accomplir cette tâche, deux types d'épreuves devaient être sélectionnés. Pour ce faire, nous avons établi une liste de critères pour chacun des deux tests. Les voici :

Premier test :

- *Objectif* : Ce test doit mesurer la performance des apprenants dans une tâche de décodage et de production de mots suffixés.
- *Critères* :
 - contexte réel de lecture
(contexte réel de lecture = contexte réel de décodage et de production de mots dérivés, l'élève ne sachant pas ce qui est évalué.)
 - codage rapide des mots suffixés
 - résultats sont en % de réussite
 - exercice rapide à effectuer
(contexte scolaire = contrainte de temps)

Deuxième test :

- *Objectif* : Ce test doit permettre d'identifier les stratégies employées par des apprenants pour décoder et produire des mots dérivés.

- *Critères :*

- l'identification des stratégies doit se faire à l'oral
- aucune contrainte de temps

À l'aide de ces critères, deux exercices ont été sélectionnés : le C-test et la tâche de rappel stimulé.

2.3.1 Le C-test

Le C-test est un instrument d'évaluation de la compétence langagière qui fut créé en 1982 par Raatz et Klein-Bradley. À cette époque, on cherchait à développer un instrument fiable et valide qui permettrait d'évaluer la compétence langagière, au même titre que le test de «closure», mais sans être aussi fastidieux à construire. C'est en se basant sur la prémisse que la langue est avant tout redondante et créative que Raatz et Klein-Bradley inventèrent le C-test. John B. Carroll (1987) résume bien leur raisonnement :

« (...) the competent user of a language suffers little loss in understanding a communication even if its redundancy is reduced, whereas the less competent user will be more handicapped. Therefore, use of a test in which redundancy is reduced should enable one to assess the competence of a language user. » (p.100)

C'est à partir de cette affirmation que les deux chercheurs créèrent le C-test, un exercice dans lequel il y a réduction de la redondance au niveau des mots, et non au niveau de la phrase, comme c'était le cas dans le test de « closure ». Sa construction est simple : on choisit tout d'abord quatre ou cinq courts textes. Pour chacun des textes, on procède de la même manière : on coupe en deux un mot sur deux et on remplace les lettres tronquées par des tirets, le nombre de tirets correspondant au nombre de lettres tronquées. Dans le cas de mots comportant un nombre impair de lettres, on applique la formule $(n + 1) / 2$, le résultat indiquant le nombre de lettres à retrancher. Finalement, afin de faciliter la compréhension

du texte, on conserve intacte la première phrase et, si on le désire, il est également possible de conserver intactes les dernières phrases d'un texte dès que l'on a atteint le nombre voulu de mots tronqués.

Cette nouvelle manière de mesurer les compétences langagières en a conquis plusieurs, à commencer par les chercheurs allemands (Grotjahn, 1994; Bauer, 1989, 1992; Stemmer, 1991, 1992; etc.), puis les chercheurs anglais (Coleman, 1994; Singleton, 1995, 1996; etc.) et finalement, son usage s'est répandu aux locuteurs de langues étrangères (Zhang, 1985; Chulasai, 1996; etc.). On a utilisé (et on utilise encore) le C-test pour mesurer les compétences langagières d'apprenants en langue première (Morris et Labelle, 2004) et en langue seconde (Dörnyei et Katona, 1992). On l'a également utilisé comme test de classement, permettant ainsi aux universités et autres institutions de classer les étudiants dans des groupes de leur niveau (Klein-Bradley, 1985). En 1991, Singleton et Little l'ont proposé comme instrument de recherche au niveau du vocabulaire en langue seconde. Cette proposition, mais surtout les conclusions auxquelles arrivèrent ces deux chercheurs, provoqua beaucoup de questionnements quant à la validité du C-test comme instrument de mesure de la connaissance et de l'organisation du vocabulaire en langue seconde (voir article de Chapelle « *Are C-tests valid measures for L2 vocabulary research?* » (1994) pour un résumé de la question). Même si, encore aujourd'hui, on continue de publier des articles pour justifier et valider l'utilisation du C-test dans un contexte de recherche (Eckes et Grotjahn, 2006), on a depuis longtemps démontré que si le C-test ne peut être utilisé pour mesurer et valider des hypothèses quant à l'organisation et l'accès au lexique mental, il peut néanmoins servir d'instrument d'évaluation du vocabulaire quand l'objectif est de mesurer certaines connaissances lexicales spécifiques telles les connaissances orthographiques et morphologiques (Chapelle, 1994, p. 179).

L'utilisation du C-test comme instrument de mesure présente en effet de nombreux avantages pour une recherche sur le traitement des mots dérivés. Tout d'abord, lecture interactive durant laquelle l'apprenant doit compléter des mots, le

C-test offre la possibilité de mesurer dans un même exercice à la fois le décodage et la production de mots suffixés. Les erreurs produites par les apprenants non natifs sont facilement identifiables puisque les locuteurs natifs obtiennent presque toujours 100% (Anckaert et Beeckmans, 1992). D'ailleurs, le fait que ce ne soit pas nécessairement des syllabes ou des morphèmes qui ont été tronqués déstabilise l'apprenant, ce qui permet d'évaluer véritablement ses capacités de décodage et de production de mots suffixés et ce, même si le texte fournit des indices à la fois sémantiques et textuels pour compléter les mots.

Un autre avantage du C-test est que celui-ci permet d'obtenir une quantité importante d'informations et ce, en peu de temps. Les données obtenues sont facilement identifiables et représentent toutes les catégories de mots (nom, préposition, verbe, adjectif, etc.). De plus, cet instrument a l'avantage de proposer un contexte réel de lecture, ce qui permet de masquer à l'apprenant l'élément ou les éléments que nous désirons véritablement évaluer, soit ici la connaissance des suffixes. Un exercice ne contenant que des mots dérivés aurait très vraisemblablement compromis nos résultats.

Malgré ses nombreux avantages, le C-test ne permet pas d'identifier les stratégies employées par ceux qui l'ont complété étant donné qu'il n'y a aucune trace écrite de ces stratégies. Les chercheurs ne peuvent qu'émettre des hypothèses quant à la nature des indices utilisés pour compléter les mots. Devant cet écueil, plusieurs chercheurs (Grotjahn, 1987; Feldmann et Stemmer, 1987; Stemmer, 1991, 1992; Little et Singleton, 1992) ont suggéré d'avoir recours à un exercice d'introspection immédiatement après le C-test afin de valider les hypothèses engendrées par celui-ci. De cette manière, il serait possible d'obtenir un portrait linguistique plus complet des apprenants. Faisant suite à cette proposition, nous avons choisi de procéder à un exercice d'introspection pour identifier les stratégies utilisées par nos participants. Mais avant de décrire cet exercice (2.2.2), regardons plus en détails l'élaboration et le contenu de nos C-tests.

2.3.1.1 Élaboration et contenu des C-tests

Dans le cadre de cette recherche, nous avons tout d'abord construit deux C-tests⁶ (voir appendice A1 et A2) que nous avons présentés aux participants à sept semaines d'intervalle. Ces C-tests, administrés en classe lors des heures de cours, nous ont permis de mesurer la performance des apprenants au niveau du décodage et de la production de mots suffixés. Les données que nous avons obtenues lors du premier C-test ont pu être validées par les résultats obtenus lors du deuxième C-test, ce dernier nous permettant de contrôler l'effet de la variable « nouveauté ».

Puisque nous voulions également identifier les stratégies utilisées par ces mêmes apprenants pour compléter les mots suffixés, nous avons créé un troisième C-test (voir appendice A3) qui allait servir d'objet de stimulation lors de notre exercice d'introspection individuelle. Ce troisième C-test a été administré sept semaines après le deuxième, en dehors des heures de cours et ce, aux participants qui étaient intéressés à poursuivre l'expérience, donc à un échantillon non aléatoire de participants.

La première étape dans l'élaboration de nos trois C-tests a été de choisir des sujets susceptibles d'être connus par tous nos participants. Étant donné la très grande diversité culturelle de ceux-ci, nous nous sommes limités à ce qu'ils avaient de commun, soit leur environnement immédiat, l'école, et leur situation actuelle, immigrant. Nous avons donc choisi de traiter des trois sujets suivants : le retour aux études, le profil de l'immigration au Québec et la gestion du temps.

Une fois nos thèmes choisis, il a fallu écrire les textes en tenant compte de plusieurs facteurs dont, entre autres, le profil culturel de nos participants ainsi que les besoins précis de la recherche. Pour ce qui est du profil culturel de nos participants, il était évident que nous devions éviter toute référence culturelle ou

⁶ En raison du contexte scolaire dans lequel notre recherche s'est déroulée, nous n'avons pas pu présenter quatre courts textes tel que le suggéraient Raatz et Klein-Bradley.

expression québécoise pouvant confondre les nouveaux immigrants. Pour ce qui est des besoins de la recherche, à savoir évaluer les capacités à décoder et à produire des mots suffixés, chaque C-test a été construit de manière à y retrouver un minimum de vingt-cinq mots suffixés tronqués, mots choisis selon la productivité de leur suffixe, nous assurant ainsi d'une variété importante de suffixes, du plus productif (*-ment*) au moins productif (*-ce*)⁷. Le choix des suffixes a été fait d'après une liste des suffixes les plus productifs, liste non publiée compilée par Longtin (2004).

Finalement, afin d'obtenir davantage d'informations sur nos participants, nous leur avons demandé d'inscrire, sur chaque C-test, leur nom et prénom, leur groupe, leur sexe ainsi que leur langue maternelle.

2.3.2 La tâche de rappel stimulé

Afin d'identifier les stratégies employées par les apprenants pour décoder et compléter des mots suffixés, il nous fallait choisir un exercice complémentaire au C-test. Plusieurs raisons nous ont amené à choisir un exercice d'introspection, entre autres, la preuve que ce type d'exercice, comme la méthode «think aloud», permet d'avoir accès à certaines stratégies de sélection du lexique (House et Blum-Kulka, 1986; Faerch et Kasper, 1987 ; Ridley et Singleton, 1995) et la démonstration de Ericsson et Simon (1984) quant à la validité des méthodes d'introspection⁸. Dans des situations similaires, plusieurs chercheurs (Haastrup, 1987; Poulisse, Bongaerts et Kellerman, 1987; Feldmann et Stemmer, 1987) ont choisi la méthode «think aloud» ainsi que la «retrospection» pour approfondir leurs résultats. Dans le cadre de notre recherche, nous avons dû écarter ces deux

⁷ Pour obtenir la liste des mots suffixés sélectionnés pour chacun des C-tests, se référer à l'appendice A.

⁸ Selon le modèle de Ericsson et Simon (1984), l'information qui vient à peine d'être traitée par l'apprenant est gardée dans la mémoire à court terme et par conséquent, est facilement accessible.

choix étant donné le profil linguistique et culturel de nos participants. Tout d'abord, la méthode « think aloud » aurait demandé une trop grande charge cognitive pour nos participants, étant donné leur niveau de français. Ensuite, les résultats que nous aurions obtenus avec la méthode de la « retrospection » auraient probablement été faussés puisque dans certaines cultures, le fait d'être repris et questionné sur les choix qui ont été faits est perçu immédiatement comme la preuve qu'il y a eu bel et bien une erreur, une faute commise. Il nous fallait donc trouver un exercice d'introspection d'une apparente neutralité qui nous permettrait d'obtenir des données de la part de participants dont le français est limité. Nous avons trouvé ce compromis dans un exercice d'introspection appelé « la tâche de rappel stimulé » (stimulated recall).

Une tâche de rappel stimulé est un instrument permettant d'explorer les pensées ou stratégies d'un apprenant au moment où il accomplissait une activité. Le souvenir des processus cognitifs est stimulé à l'aide d'un rappel, visuel ou oral, et de questions neutres auxquelles l'apprenant doit répondre à haute voix. Les réponses sont enregistrées. Il est à noter que pour qu'une tâche de rappel stimulé soit efficace (à 95%), elle doit généralement être administrée à l'intérieur de 48 heures suivant l'événement (Bloom, 1954). Néanmoins, la plupart des chercheurs s'accordent pour dire que pour avoir accès aux processus cognitifs lors d'une tâche de rappel stimulé, il faut que celle-ci ait lieu immédiatement après l'événement en question (Ericsson et Simon, 1987).

Tout comme le C-test, la tâche de rappel stimulé comme instrument d'évaluation présente plusieurs avantages. Dans leur livre « *Stimulated Recall Methodology in Second Language Research* », Susan M. Gass et Alison Mackey (2000) en font l'énumération :

- Outil qui permet d'identifier les processus cognitifs qui n'auraient pu être observés autrement.
- Outil qui permet d'isoler un événement en particulier d'un ensemble d'éléments ou d'activités.

- Outil qui permet d'identifier le type de connaissances que le sujet a utilisé pour résoudre un problème, prendre une décision linguistique ou accomplir une tâche de production et/ou de compréhension.
- Outil qui permet d'observer comment cette connaissance a été organisée (dans le lexique mental par exemple).
- Outil qui permet de vérifier si un processus cognitif en particulier, telle une prise de décision, a été utilisé et à quel moment il a été utilisé. (p.21)

De plus, dans leurs conclusions, Cohen (1996, 1998) et Matsumoto (1993) mentionnent que la tâche de rappel stimulé est un outil utile lorsque l'on fait de la recherche au niveau des stratégies. Pour ce qui est de valider la procédure, soit une tâche de type entrevue, Rod Ellis (1997) soutient que les apprenants en général sont conscients des stratégies qu'ils utilisent et, quand on le leur demande, peuvent expliquer ce qu'ils ont fait pour compléter des exercices.

Malgré ses nombreux avantages, la tâche de rappel stimulé a également ses écueils, surtout lorsqu'elle est utilisée en langue seconde. Tout d'abord, si le niveau de français des sujets n'est pas assez bon, ils risquent de ne pas être capables a) de répondre aux questions et b) d'identifier dans des mots précis les stratégies qu'ils ont utilisées. La réussite de la tâche de rappel stimulé est donc dépendante, en quelque sorte, des habiletés verbales des participants. De plus, il est possible que certains apprenants modifient légèrement leurs réponses dans le but de plaire à l'interviewer. Par conséquent, afin d'éviter de se retrouver dans l'une ou l'autre de ces situations, nous avons spécifiquement choisi a) d'évaluer des étudiants adultes du niveau avancé du certificat en français écrit car avant d'arriver à ce niveau, ces étudiants ont dû satisfaire à des exigences orales; b) de bâtir un protocole de recherche qui limite l'interférence possible de l'interviewer et met en confiance le participant en lui précisant le but de l'entrevue et en l'assurant qu'il n'y a aucune mauvaise réponse, que toutes les réponses sont bonnes.

2.3.2.1 Élaboration et contenu de la tâche de rappel stimulé

La tâche de rappel stimulé a été administrée individuellement et ce, immédiatement après le troisième C-test, ce dernier servant d'objet de rappel lors de l'entrevue. Afin d'être en mesure d'obtenir des données pertinentes pour notre recherche, nous avons d'abord sélectionné les mots dérivés du C-test pour lesquels nous voulions connaître les stratégies de décodage et de production. Nous avons choisi des mots dérivés représentatifs de toutes les catégories de dérivation suffixale, du verbe au nom, de l'adjectif au nom, etc. Nous avons ensuite bâti un protocole de recherche selon le modèle proposé par Gass et Mackey. Dans ce protocole, nous avons établi les questions qui allaient être posées, les mots d'encouragement que pouvait dire l'interviewer et les procédures à suivre pour que toutes les entrevues soient similaires et qu'il n'y ait aucune interférence de la part de l'interviewer (voir protocole, appendice B1).

Nos deux instruments de mesure ayant été décrits, passons maintenant au déroulement de notre collecte de données.

2.4 Déroulement et collecte des données

Dans un premier temps, nous avons fait l'essai de nos instruments de mesure auprès de cinq adultes parlant le français comme langue seconde ainsi que de quatre adultes locuteurs natifs du français. Ces essais ont eu lieu au début du mois de septembre 2004 et ils nous ont permis d'ajuster nos instruments et de contrôler la variable « temps ».

Dans un deuxième temps, nous avons contacté la responsable du niveau avancé du « *Certificat en français écrit pour non-francophones* » pour lui présenter notre projet (voir appendice C1). Une fois son autorisation accordée, nous avons rencontré les professeurs intéressés. Lors de cette rencontre, qui a eu lieu en

septembre, nous leur avons expliqué le projet et nous avons ensuite discuté d'un échéancier. Les deux premières épreuves, soit les deux premiers C-tests, allaient avoir lieu en classe, pendant les heures de cours, en septembre et en décembre. Une période de 40 minutes a été prévue pour le premier C-test et une période de 25 minutes a été prévue pour le deuxième C-test.

Il est également important de mentionner que, entre nos deux épreuves, nos sujets n'ont pas étudié les suffixes en classe. Dans deux groupes sur trois, l'enseignement a porté sur les pronoms personnels, possessifs, relatifs et démonstratifs; les verbes transitifs directs et indirects; les prépositions ainsi que les participes passés. Quant au troisième groupe, l'enseignement a porté sur la compréhension de discours de type universitaire et de textes provenant de diverses disciplines, la prise de notes ainsi que la lettre de plainte et l'essai. Afin de s'assurer qu'il n'y aurait aucun enseignement systématique des suffixes, nous en avons également discuté avec les professeurs concernés. Ces derniers nous ont promis qu'ils n'y auraient pas recours.

À la fin du mois de septembre, lors de la première collecte de données, nous avons d'abord présenté notre recherche aux étudiants et nous leur avons ensuite demandé de signer un formulaire de consentement nous autorisant à utiliser leurs résultats. Afin de nous assurer que tous comprenaient bien les enjeux, les formulaires de consentement avaient été traduits dans quatre langues en plus du français : l'anglais, le russe, l'espagnol et le mandarin.

Après avoir recueilli les formulaires signés, nous leur avons distribué le premier C-test. Une fois que celui-ci a été brièvement expliqué, nous nous sommes assurés que tous les étudiants avaient bien inscrit en haut de leur feuille leur nom, prénom, groupe, sexe ainsi que leur langue maternelle. Les étudiants ont ensuite eu 15 minutes pour compléter le C-test. Cette première séance s'est terminée par la remise du C-test complété.

Nous avons procédé à la deuxième collecte de données dans les mêmes groupes à la mi-décembre, soit à la toute fin de la session d'automne 2004. Cette séance s'est déroulée exactement comme la première collecte à l'exception près de l'explication de la recherche et de la signature des formulaires de consentement. À la fin de cette deuxième collecte, nous avons demandé à tous les étudiants intéressés à poursuivre l'expérience de bien vouloir nous laisser leurs coordonnées. La troisième collecte de données allait être effectuée sur une base volontaire étant donné qu'il aurait été difficile de rejoindre les mêmes étudiants à la session d'hiver, ceux-ci ne s'étant pas nécessairement inscrits aux même cours.

Au début du mois de janvier 2005, nous avons contacté tous les étudiants qui nous avaient laissé leurs coordonnées et avons donné à chacun un rendez-vous pour la fin du mois de janvier, à raison d'une rencontre par heure. Lors de cette troisième collecte de données, qui s'est déroulée dans un local de l'université, nous avons demandé à chacun des participants de compléter un autre C-test (d'une durée de 15 minutes) et leur avons ensuite proposé de poursuivre la rencontre par une entrevue qui allait être enregistrée et durant laquelle nous allions leur poser quelques questions sur l'exercice qu'ils venaient de faire. Sur un total de vingt-quatre entrevues, deux n'ont pas pu être enregistrées à cause d'un problème technique avec l'enregistreuse. Néanmoins, les réponses des sujets ainsi que les questions de l'interviewer ont été soigneusement notées.

La durée des entrevues de rappel stimulé a varié entre 15 et 30 minutes. Par la suite, toutes les questions et réponses des entrevues ont été transférées sur des disquettes afin d'être ultérieurement analysées. La manière dont nous avons procédé pour coder nos résultats ainsi que la description des analyses que nous avons effectuées seront présentées dans la prochaine section.

2.5 Codage et analyse des données

Dans cette section, nous décrivons les outils que nous avons utilisés pour coder et analyser les données que nous avons recueillies. Il sera tout d'abord question du codage des C-tests (2.5.1) et des réponses obtenues lors de la tâche de rappel stimulé (2.5.2). Ensuite, nous présenterons brièvement les différentes analyses que nous avons choisi d'effectuer pour analyser nos résultats.

2.5.1 Codage – C-tests

Dans un premier temps, nous avons fait un tri des C-tests afin de ne garder que ceux qui avaient été entièrement complétés. Les C-tests étaient rejetés s'ils répondaient à l'un ou à l'autre de ces deux critères :

- a) moins de 60 % des mots ont été complétés
- b) le participant ne semble pas avoir eu le temps de terminer l'exercice; les mots du dernier paragraphe en entier n'ont pas pu être complétés.

Le nombre de C-tests rejetés pour chacune des épreuves sera décrit en détails au chapitre III.

Dans un deuxième temps, toutes les réponses des C-tests acceptés ont été compilées dans le logiciel EXCEL. Elles ont ensuite été codées selon la grille suivante :

Tableau 2.3
Grille de codage des C-tests

0	=	aucun mot
1	=	un mot, mais il n'est pas français
2	=	un mot français, mais ce n'est pas le bon
3	=	un mot français, mais erreur de suffixe
4	=	bon mot, mais erreur d'orthographe
5	=	bon mot, mais mauvaise flexion
6	=	bon mot écrit correctement

Si l'étudiant n'avait tout simplement pas complété le mot, ce mot était codé 0. Si la réponse de l'étudiant ne correspondait pas du tout à un mot français, soit qu'il n'existait pas ou soit qu'il correspondait à un mot appartenant à une autre langue, ce mot était codé 1. Si le participant avait bien écrit un mot français, mais que ce n'était pas le bon mot, ce mot était codé 2. Si le mot français choisi n'était pas le bon à cause d'une erreur de suffixe, ce mot était codé 3. Si le participant avait écrit le bon mot mais avait fait une erreur au niveau de l'orthographe, ce mot était codé 4. Le bon mot mais comportant une erreur de flexion équivalait à un 5. Finalement, si le mot complété était le bon et qu'il avait été correctement orthographié, ce mot était codé 6.

Ce codage nous a permis, tout d'abord, d'établir en pourcentage (%) le taux global de réussite de chaque C-test pour chaque participant. Nous avons également pu déterminer le taux de réussite des mots suffixés. De plus, nous avons pu identifier les mots les plus réussis, ceux les moins réussis et ce, non seulement pour l'ensemble de notre population, mais également pour chacun de nos deux groupes à l'étude (LPA et LBA).

2.5.2 Codage - Tâche de rappel stimulé

Chaque entrevue a été enregistrée, puis transférée dans un logiciel permettant l'écoute des enregistrements via un ordinateur. Chaque entrevue a été écoutée deux fois et les réponses ont été codées selon la grille de codage suivante :

Tableau 2.4
Grille de codage des stratégies

Catégorie	Stratégies
Contexte élargi	<input type="checkbox"/> Sens global du texte <input type="checkbox"/> Sens du paragraphe <input type="checkbox"/> Sens de la phrase <input type="checkbox"/> Mots-clés du texte
Contexte immédiat	<input type="checkbox"/> Mot avant <input type="checkbox"/> Mot après <input type="checkbox"/> Expression <input type="checkbox"/> Les premières lettres données <input type="checkbox"/> Les tirets
Analyse grammaticale et lexicale	<input type="checkbox"/> La nature du mot à compléter <input type="checkbox"/> La fonction du mot à compléter <input type="checkbox"/> La morphologie flexionnelle <input type="checkbox"/> La morphologie dérivationnelle <input type="checkbox"/> Les règles de grammaire <input type="checkbox"/> Recours au lexique et au champ lexical
Autres	<input type="checkbox"/> Ne sait pas

Cette grille d'observation a été inspirée par les différentes listes de stratégies proposées en langue seconde (voir le point 1.2.1) et a été validée lors des tests pilotes.

La catégorie « Contexte élargi » comprend toutes les stratégies qui font référence au sens. Le participant peut s'être fié au sens du texte en général, du paragraphe ou même au sens de la phrase. Il peut également s'être inspiré des autres mots dans le texte, comme les mots du titre par exemple, pour compléter un mot tronqué.

La catégorie « Contexte immédiat » fait référence à l'environnement immédiat du mot : le mot avant et le mot après; un groupe de mots qui forme une expression et qui comprend le mot tronqué; les trois premières lettres écrites du mot ou même le nombre de tirets.

La catégorie « Analyse grammaticale et lexicale » comprend toutes les stratégies qui ont un lien avec la grammaire du français ainsi que le vocabulaire connu. Pour compléter les mots tronqués, le participant peut s'être fié à la nature du mot à compléter ou même à sa fonction. Il peut avoir fait référence à la morphologie flexionnelle en parlant des règles d'accord ou même à la morphologie dérivationnelle en mentionnant le suffixe. Il peut avoir cherché dans son lexique mental tous les mots qu'ils connaissaient faisant partis du même champ lexical (ex. : *difficile*, *difficilement*, *difficulté*) ou de la même famille de mots ou même s'être souvenu d'une règle quelconque, bonne ou erronée, apprise en classe.

Finalement, même s'il ne s'agit pas d'une stratégie, nous avons voulu compiler tous les « Je-ne-sais-pas ». Cette réponse, en soi, en dit déjà beaucoup. Nous aborderons ce sujet lors de la discussion de nos résultats (chapitre III).

Puisque plus d'une stratégie pouvait être nommée pour un mot, les stratégies nommées ont été identifiées selon leur ordre d'apparition. Une fois toutes les stratégies identifiées, elles ont été compilées dans EXCEL.

Cette grille de codage nous a permis d'identifier les catégories de stratégies les plus utilisées, celles les moins utilisées et ce, non seulement pour l'ensemble de notre population, mais également pour chacun de nos deux groupes (LPA et LBA).

2.5.3 Analyses

Avant de présenter nos résultats, voici brièvement les types d'analyses qui ont été effectués pour chacun des exercices.

a) Analyses – C-test

Comme nous l'avons mentionné précédemment (2.5.1), les résultats des C-tests ont d'abord été compilés dans le logiciel EXCEL. Ce logiciel nous a permis de faire des moyennes et de comparer plusieurs choses :

- les mots les plus réussis dans les deux groupes
- les mots les moins réussis dans les deux groupes
- les mots qui différencient les deux groupes
- les taux de réussite (global et mots dérivés) des deux groupes

Une fois ces comparaisons effectuées, les taux de réussite ont été soumis à des tests statistiques avec le logiciel *Statistical Package for the Social Sciences* (SPSS) afin de savoir si la L1 avait un impact significatif sur ces taux. Pour ce faire, nous avons effectué des analyses ANOVA. La langue maternelle devenait un facteur d'influence si la différence significative était ≤ 0.05 .

Nous avons ensuite comparé les résultats des analyses ANOVA pour les trois C-tests. Ceci nous a permis de contrôler la variable nouveauté, mais surtout de vérifier s'il y avait eu une certaine amélioration significative, c'est-à-dire s'il y avait eu un certain apprentissage dans un groupe ou dans l'autre ou s'il existait une différence significative réelle et persistante entre les deux groupes.

b) Analyses – Tâche de rappel stimulé

Pour ce qui est de la tâche de rappel stimulé, nous avons d'abord entrepris de compiler les résultats obtenus dans le logiciel EXCEL afin d'être en mesure de comparer, quantitativement, les différentes stratégies employées. Ceci nous a permis d'identifier la catégorie de stratégies la plus nommée et celle la moins mentionnée; la stratégie la plus populaire et celle qui l'a moins été; etc.

Finalement, nous avons cherché à savoir s'il existait une correspondance entre les stratégies qui avaient été nommées et les réponses écrites lors du troisième C-test. Pour y parvenir, nous avons effectué un nouveau codage à l'aide de la grille d'évaluation suivante :

Tableau 2.5
Grille de codage : correspondance entre stratégie et réponse

Bonne stratégie et bonne réponse	1
Bonne stratégie et mauvaise réponse	2
Mauvaise stratégie et bonne réponse	3
Mauvaise stratégie et mauvaise réponse	4

Nous avons codé à nouveau tous les mots du troisième C-test qui avaient été ciblés lors de la tâche de rappel stimulé. Dans ce dessein, nous avons comparé chaque réponse écrite par les participants avec la ou les stratégie(s) qu'ils avaient identifiée(s) lors de l'entrevue. Si la stratégie était bonne et qu'ils avaient effectivement écrit le bon mot, ce mot était codé 1. Si la stratégie était bonne, mais qu'ils avaient écrit la mauvaise réponse, le mot était codé 2. S'ils avaient mentionné une stratégie qui ne pouvait pas aider à compléter le mot, mais qu'ils l'avaient tout de même bien écrit, ce mot était codé 3. Finalement, s'ils avaient mentionné une stratégie inutile et qu'ils n'avaient pas réussi à trouver le bon mot, celui-ci était codé 4.

Ce nouveau codage nous a permis, encore une fois, de comparer nos deux groupes à l'étude et de vérifier si la correspondance entre la stratégie et la réponse était la même pour les deux groupes. L'explication de nos codages et de nos analyses étant terminée, passons maintenant à la présentation des résultats.

CHAPITRE III

PRÉSENTATION ET DISCUSSION DES RÉSULTATS

Dans le présent chapitre, il sera question des résultats que nous avons obtenus lors de nos épreuves et de la signification de ceux-ci par rapport à nos hypothèses de recherche. Dans un premier temps, nous présenterons les résultats de nos C-tests (3.1). Ces résultats seront d'abord précédés par la description de la répartition de nos participants ainsi que des problèmes que nous avons rencontrés lors du codage des données.

Dans un deuxième temps, nous nous intéresserons aux réponses données par nos participants lors de la tâche de rappel stimulé (3.2). Tout comme pour les C-tests, nous débuterons cette section par une description de la répartition de nos sujets. Nous enchaînerons ensuite avec les problèmes rencontrés lors de l'administration de la tâche de rappel stimulé ainsi que ceux survenus lors du codage. Nous terminerons par la présentation des données recueillies et tenterons de faire des liens entre ces résultats et ceux obtenus lors du troisième C-test qui, rappelons-le, a servi d'objet de rappel pour l'exercice.

Finalement, nous conclurons ce chapitre en faisant la synthèse de nos observations et en vérifiant si celles-ci nous permettent de confirmer nos deux hypothèses de recherche (3.3).

3.1 Résultats – C-tests

Nous tenons d'abord à préciser que nous ne présenterons pas les résultats des trois C-tests dans la même section. Nous discuterons d'abord des résultats obtenus lors du premier et du deuxième C-test (3.1.1) et, ensuite, ceux recueillis lors du troisième C-test (3.1.2). Deux arguments majeurs nous ont amené à faire ce

choix. Tout d'abord, notre deuxième C-test servait avant tout à contrôler la variable « nouveauté », qui a bien souvent un impact non négligeable sur les résultats, surtout lorsque les participants ne sont pas familiers avec le type d'exercice à effectuer. En présentant conjointement les résultats du premier et du deuxième C-test, nous nous assurons d'être en mesure d'identifier cet impact et de l'écarter. Ensuite, si nous avons choisi de présenter les résultats du 3^e C-test séparément, c'est que cette épreuve a été créée, avant tout, afin de servir d'objet de rappel pour notre tâche de rappel stimulé. Nous avons choisi d'en présenter les résultats dans cette section, et non dans la section concernant la tâche de rappel stimulé, car nous croyons que ceux-ci pourront être utilisés afin d'appuyer (ou de nuancer) les observations que nous aurons faites lors des deux premiers C-tests.

3.1.1 Résultats – C-test 1 et C-test 2

Avant de présenter la répartition de nos participants, rappelons le contexte dans lequel se sont déroulés ces deux épreuves. Notre recherche a été menée auprès d'étudiants de niveau universitaire suivant des cours de français écrit dans le but d'obtenir un certificat en français écrit pour non francophones. Nos deux C-tests ont été administrés à ces étudiants durant les heures de cours et ce, à sept semaines d'intervalle. Tous les étudiants pouvaient refuser de participer à la recherche et avaient la possibilité, à tout moment, d'annuler leur participation. Chaque groupe de participants a reçu la même consigne pour compléter les C-tests et s'est vu allouer le même délai, soit 15 minutes, pour les compléter.

3.1.1.1 Répartition des participants

a) C-tests rejetés

Comme nous l'avons mentionné précédemment (voir le point 2.3), nous avons évalué des étudiants de niveau avancé inscrits à un certificat en français écrit pour

non-francophones. Au total, près d'une centaine d'étudiants ont complété les C-tests. Malheureusement, plusieurs C-tests ont dû être rejetés dès le départ pour l'une ou l'autre des raisons suivantes :

- a) L'étudiant a été absent une journée et n'a donc pas pu compléter les deux C-tests.
- b) L'étudiant n'a pas réussi à terminer l'exercice dans le délai requis⁹.
- c) L'étudiant n'a pas réussi à compléter plus de 50 % des mots tronqués.

Par conséquent, de la centaine d'étudiants que nous avions au départ, seulement 63 d'entre eux ont complété en bonne et due forme les deux C-tests. Nous avons regroupé les cas d'exclusion que nous vous présentons dans le tableau suivant :

Tableau 3.1
Nombre de C-tests rejetés selon le type d'explication

Raison	Raison a : Absence du participant à l'une ou l'autre des épreuves	Raison b : C-test non complété dans le délai requis	Raison c : C-test non complété (moins de 50% des mots tronqués complétés)	Total
<i>Total</i>	19	4	5	28

D'un premier coup d'œil, nous pouvons constater que la principale raison pour laquelle nous avons dû rejeter des C-tests a été l'absence des participants à l'une ou à l'autre des épreuves. Curieusement, cette absence, dans 84% des cas, s'est produite lors du deuxième C-test. Ces absences peuvent être justifiées par de nombreux facteurs dont, entre autres :

- Les épreuves ont été administrées au tout début du cours. Ces élèves se sont absentés ces jours-là pour des raisons personnelles ou sont tout simplement

⁹ Afin de s'assurer que le participant avait bel et bien manqué de temps pour terminer l'exercice, et non pas eu de la difficulté à compléter 50 % des mots tronqués, nous nous sommes basés sur le critère suivant : le participant a manqué de temps s'il a tenté de répondre à la majorité des mots du texte et que seuls les mots tronqués du dernier paragraphe n'ont pas été complétés.

arrivés en retard au cours et n'ont pas eu la possibilité de compléter les exercices.

- Les épreuves n'étant pas obligatoires pour le cours (aucun crédit ne leur était accordé s'ils complétaient les C-tests) et la date du deuxième C-test ayant été annoncée à l'avance en classe, certains étudiants ont pu tout simplement décider de se présenter au cours une fois l'épreuve terminée (étant donné que les C-tests ont tous été administrés au début du cours).
- Les participants ayant complété le premier C-test et l'ayant trouvé difficile ont choisi de ne pas se présenter pour le deuxième.

b) Participants

La répartition de nos participants pour ces deux premières épreuves sera présentée selon leur sexe et le groupe auquel ils appartiennent, soit le groupe dont la langue maternelle écrite contient peu d'affixes (LPA) ou le groupe dont la langue maternelle écrite contient beaucoup d'affixes (LBA).

Tableau 3.2.

Répartition des participants selon le sexe et la L1 – C-test 1 et 2

Profil linguistique	Total d'étudiants	
	<i>M</i>	<i>F</i>
LPA	6	30
LBA	8	19
Total	63	

À la lecture de ce tableau, on constate que la majorité des participants à notre étude était des femmes (78%). Par ailleurs, plus d'étudiants appartenant au groupe LPA (57%) ont participé à notre recherche que d'étudiants appartenant au groupe LBA (43%). La différence demeure toutefois minime.

3.1.1.2 Problèmes encourus lors du codage des C-tests

Coder des données n'est pas chose facile, surtout lorsque les données recueillies consistent en de courtes réponses manuscrites. Non seulement faut-il être en mesure de comprendre la calligraphie de chaque participant, mais il faut également résoudre certains problèmes générés par le type d'exercice.

Dans le cadre du codage des C-tests, nous avons rencontré quelques problèmes mineurs que nous présenterons dans la présente section. Leur ordre de présentation est tout à fait aléatoire.

a) Omission des accents : erreur d'orthographe ?

Plusieurs mots en français comportent des accents. Devait-on considérer l'omission de ces accents comme des erreurs d'orthographe ? Deux raisons nous ont amené à ne pas considérer l'absence de certains accents comme des erreurs. Tout d'abord, la récente proposition de la réforme de l'orthographe, instiguée par le groupe RENOUVO (2005), suggère, entre autres, l'abolition de l'accent circonflexe sur les lettres *i* et *u*, le changement du *è* pour un *é* devant un *e* muet et le déplacement du tréma sur la lettre *u*. Bien que ces modifications visent à supprimer des anomalies de l'orthographe française, les graphies anciennes restent admises, ce qui venait compliquer notre codage. De plus, le but de notre recherche étant de vérifier si les participants étaient capables de compléter des mots tronqués, et non pas nécessairement de les écrire correctement avec les accents, nous avons décidé de ne pas tenir compte de l'omission des accents sauf dans les deux cas suivants :

- Lorsque l'accent détermine la catégorie grammaticale du mot.

Ex. : « a » : le verbe *avoir* conjugué à la 3^e personne du singulier de l'indicatif présent
 « à » : préposition

- Lorsque l'accent vient préciser le temps de verbe.

Ex. : « confronte » : le verbe *confronter* conjugué à la 1^{re} ou 3^e personne du singulier du présent de l'indicatif

« confronté » : le verbe *confronter* au participe passé

b) Coquille ou erreur d'orthographe ?

Les coquilles sont des erreurs courantes lorsque la tâche écrite doit être accomplie dans un délai précis. Voici un exemple de ce type d'erreur extrait du 2^e C-test :

Ex. : Participant 111 Mot à compléter : niv__ _ (niveau)
 Mot complété : *nivaeu*

La distinction entre la coquille et la véritable erreur devient cependant moins claire lorsque nous nous retrouvons avec un exemple comme celui-ci (extrait du 1^{er} C-test):

Ex. : Participant 112 Mot à compléter : effec__ _ (effectuer)
 Mot complété : *effecteur*

A-t-on affaire ici à une simple coquille ou à une erreur de dérivation ? Est-ce que le participant savait que le mot à compléter était un verbe ou a-t-il tout simplement cherché un suffixe pour former un nom ? Étant donné que nous n'avions aucun indice textuel nous permettant de répondre à ces questions, et pour être juste avec tous les participants, nous avons tout simplement décidé de considérer toute coquille comme une erreur.

c) Deux réponses au lieu d'une

Il est souvent arrivé que des participants aient d'abord complété le mot en se servant des tirets et aient ensuite trouvé un deuxième mot qu'ils ont écrit au-dessus du premier. Dans de pareils cas, nous n'avons pas tenu compte du deuxième mot. Seules les lettres écrites sur les tirets ont été prises en considération.

d) Répétition de la même lettre : erreur d'orthographe ?

Lors du codage, nous avons pu observer dans de nombreuses occasions une répétition de la dernière lettre donnée. Illustrons ce type d'erreur à l'aide d'un exemple extrait du 1^{er} C-test:

Ex. : Participant 105

Mot à compléter : déroul_ _ _ _ _ (déroulement)

Mot complété : *déroullemen* (répétition du « l »)

Devait-on considérer cette répétition comme une erreur d'orthographe ? D'un commun accord, nous avons décidé de ne pas tenir compte de cette répétition lors de la correction et ce, pour deux raisons :

- 1- Le C-test est un exercice demandant beaucoup de concentration; il y avait donc un haut risque de commettre ce type d'erreur.
- 2- Les tirets dans nos trois C-tests étaient assez petits (le premier tiret étant d'ailleurs littéralement collé à la dernière lettre donnée du mot à compléter).

Si nous n'en avons pas tenu compte lors du codage, nous avons néanmoins constaté que cette répétition provoquait systématiquement une erreur quand le mot à compléter devait être écrit au pluriel. La lettre répétée occupant un tiret, il ne restait plus de tiret pour l'ajout du « s » ou du « x ».

e) Erreur de flexion ou erreur de dérivation ?

Étant donné que nous avons bâti nos C-tests afin d'y retrouver plusieurs mots dérivés tronqués, nous nous attendions à observer un nombre important d'erreurs au niveau des suffixes choisis. Or, ce que nous avons constaté lors du codage, c'est que lorsque le suffixe choisi n'était pas le bon, il était souvent accompagné d'une erreur de flexion. Voici un exemple extrait du 2^e C-test illustrant le type de problème rencontré :

Ex. : Participant 111 Mot à compléter : scola_ _ _ _ (scolarité)
 Mot complété : *scolaires*

Dans cet exemple, nous retrouvons à la fois une erreur de dérivation, c'est-à-dire le mauvais choix du suffixe, mais également une erreur de flexion, ici l'ajout d'un « s » alors que le mot à compléter devait être écrit au singulier. Comment alors procéder au codage de cette réponse ? Pour résoudre ce dilemme, nous nous sommes basés sur les indices textuels afin de déterminer l'ordre d'importance des erreurs. Dans ce cas-ci, nous avons jugé qu'il y avait avant tout une erreur de dérivation. Le mot à compléter devait être un nom car il avait pour fonction de compléter le mot « niveau » et le choix d'un suffixe adjectival démontre clairement que le participant ne s'est pas fié au sens du contexte immédiat.

3.1.1.3 Résultats - Réussite globale

Dans cette section, nous présenterons les résultats globaux de nos deux premiers C-tests. Nous les avons regroupés en tableaux afin d'en faciliter la compréhension. Ces tableaux, au nombre de deux, nous permettront de mesurer l'influence de deux variables : le sexe des participants (tableau 3.3) et les caractéristiques morphologiques du système d'écriture de leur L1 (tableau 3.4).

a) Réussite globale selon le sexe

Notre objectif étant d'isoler toutes les variables possibles, nous avons voulu vérifier, dans un premier temps, si la variable « sexe » avait eu un impact sur les performances de nos participants.

Tableau 3.3
Réussite globale selon le sexe – C-tests 1 et 2

Sexe	C-test 1 Moyenne	C-test 2 Moyenne
Masculin	67 %	67 %
Féminin	59 %	65 %

Ce que nous pouvons observer dans ce tableau, c'est qu'il y a tout d'abord très peu de différences au niveau du pourcentage de réussite entre les participants de sexe masculin et ceux de sexe féminin. Si la différence entre les moyennes est plus grande lors de la première épreuve (8%), elle diminue lors de la deuxième épreuve (2 %), les participants de sexe féminin ayant obtenu de meilleurs résultats. D'après des analyses ANOVA, ces différences ne sont pas statistiquement significatives (C-test 1 : Sig. 0,139; C-test 2 : Sig. 0,578), ce qui confirme qu'il n'existe pas de différence significative au niveau de la performance entre les hommes et les femmes de notre échantillon. Par conséquent, pour la suite de nos analyses, nous ne tiendrons plus compte de la variable « sexe ».

b) Réussite globale selon les caractéristiques morphologiques du système d'écriture de la L1 des participants

L'une de nos hypothèses de recherche étant d'affirmer qu'un participant ayant une L1 écrite riche en affixes (LBA) aura de meilleurs résultats dans une tâche de décodage et de production de mots dérivés en français langue seconde qu'un participant ayant une L1 pauvre en affixes (LPA), nous avons pensé qu'il serait intéressant de vérifier cette hypothèse au niveau de la réussite globale du C-test. Est-ce que les participants du groupe LBA ont obtenu une meilleure moyenne lors de l'épreuve du C-test que les participants du groupe LPA ? Afin de répondre à cette question, nous avons d'abord réparti nos participants en deux groupes selon les caractéristiques morphologiques du système d'écriture de leur langue première et avons ensuite calculé, à l'aide de leurs résultats, les moyennes obtenues lors du premier et du deuxième C-test. Voici les résultats que nous avons obtenus :

Tableau 3.4
Réussite globale selon les caractéristiques morphologiques
du système d'écriture de la L1 - C-tests 1 et 2

Richesse morpho de la L1	C-test 1 Moyenne	ANOVA C-test 1 Différence entre LBA et LPA			C-test 2 Moyenne	ANOVA C-test 2 Différence entre LBA et LPA		
		Df	F	Sig.		Df	F	Sig.
LBA	72 %	1	27,728	0,000*	70 %	1	7,154	0,01*
LPA	53 %				62 %			

* significatif $\leq 0,05$

Comme nous pouvons le constater, alors que la moyenne obtenue par le groupe LBA est restée sensiblement la même (-2%) lors des deux épreuves, la moyenne obtenue par le groupe LPA lors du 2^e C-test a fait un bond de 8%. Selon nous, il est possible que ce soit le résultat de la familiarité des participants avec le type de tâche demandé, bien que rien ne le prouve. D'un autre côté, malgré cette amélioration, la différence entre les moyennes de ces deux groupes demeure importante (19% pour le 1^{er} C-test et 8% pour le 2^e C-test). Est-ce que ces différences sont statistiquement significatives ? Pouvons-nous affirmer que tous les apprenants ayant une L1 écrite pauvre en affixes réussissent moins bien à un C-test en français langue seconde que des apprenants ayant une L1 riche en affixes et ce, à niveau égal ? Des analyses ANOVA nous permettent de conclure que les différences observées entre les moyennes des deux groupes sont statistiquement significatives. Nous pouvons donc affirmer que des apprenants appartenant au groupe LBA obtiennent de meilleurs résultats dans un C-test en français langue seconde que des apprenants du groupe LPA. Ceci vient corroborer les résultats obtenus dans des recherches similaires pour l'anglais langue seconde (Muljani et al, 1998; Koda, 1999).

Si nous avons trouvé des différences significatives au niveau de la performance globale, en arriverons-nous à la même conclusion pour ce qui est du décodage et

de la production de mots dérivés ? C'est à cette question que nous tenterons de répondre dans la prochaine section.

3.1.1.4 Résultats - Réussite des mots dérivés

Les mots dérivés étant la pierre angulaire de notre recherche, nous en avons intégré plusieurs dans chaque épreuve. On en dénombre respectivement 26 et 32 dans le premier et le deuxième C-test. Notre choix des mots dérivés a été basé sur la fréquence et la productivité des suffixes (pour de plus amples détails sur la construction des C-tests, se référer à la section 2.2.1.1).

Dans cette section, nous mesurerons l'influence des caractéristiques morphologiques du système d'écriture de la langue première des participants, une variable directement en lien avec notre première hypothèse de recherche. Celle-ci, rappelons-le, était de vérifier si des apprenants du français langue seconde ayant une L1 écrite riche en affixes (LBA) allaient obtenir de meilleurs résultats dans une tâche de décodage et de production de mots dérivés que des apprenants ayant une L1 pauvre en affixes (LPA). La littérature portant sur le sujet confirme cette hypothèse (voir chapitre I).

Afin de vérifier à notre tour la véracité de cette hypothèse, nous avons refait des calculs en répartissant en deux groupes nos participants (LBA et LPA) et en calculant la moyenne des deux groupes pour la réussite des mots dérivés de chaque C-test. Voici les résultats que nous avons obtenus :

Tableau 3.5
Réussite des mots dérivés selon les caractéristiques morphologiques
du système d'écriture de la L1 - C-tests 1 et 2

Richesse morpho de la L1	C-test 1 Moyenne	ANOVA C-test 1 Différence entre LBA et LPA			C-test 2 Moyenne	ANOVA C-test 2 Différence entre LBA et LPA		
		Df	F	Sig.		Df	F	Sig.
LBA	75 %	1	20,723	0,000*	62 %	1	12,264	0,001*
LPA	52 %				47 %			

* significatif $\leq 0,005$

Tout d'abord, nous pouvons observer, pour les deux groupes, une diminution du pourcentage de réussite des mots dérivés au 2^e C-test. Nous expliquons cette diminution non pas par le choix des suffixes, ceux-ci étant sensiblement les mêmes dans les deux épreuves, mais plutôt par la fréquence de ces mots dérivés à l'écrit. En effet, alors que l'ensemble des mots dérivés du premier C-test et du deuxième C-test ont la même fréquence dans la langue orale (1^{er} C-test : taux de fréquence films (tff) : 11,86; 2^e C-test : tff: 11,52¹⁰), un écart assez grand sépare les deux groupes de mots en ce qui a trait à leur fréquence dans la langue écrite (1^{er} C-test : taux de fréquence livres (tfl) : 22,29; 2^e C-test : tfl : 15,98). Nous nous permettons de supposer que la diminution du pourcentage de réussite des mots dérivés du deuxième C-test est causée par une présence beaucoup moins importante de ces mots dans la langue écrite comparativement aux mots dérivés du premier C-test.

Ensuite, nous pouvons constater que le groupe LBA a mieux performé au niveau des mots dérivés que le groupe LPA et ce, même si sa moyenne a diminué lors du 2^e C-test. Avec des différences de l'ordre de 23% au premier C-test et de 15% au

¹⁰ Ces taux ont été tirés du site « www.lexique.org », subventionné par le CNRS. Les taux de fréquence présentés ici ont été calculés selon la formule suivante : la somme des fréquences du lemme et de ses formes fléchies sur 1 million d'occurrences. Pour de plus amples renseignements sur le calcul de ces taux de fréquence, se référer au site www.lexique.org

deuxième C-test, le groupe LBA semble de loin supérieur dans la maîtrise des mots dérivés. Cette supériorité est confirmée par les analyses ANOVA qui nous permettent d'affirmer que la différence entre les moyennes est statistiquement significative dans les deux C-tests. Ceci signifie que statistiquement, les apprenants du groupe LBA réussissent mieux à une tâche de décodage et de production de mots suffixés en français langue seconde que les apprenants du groupe LPA. Encore une fois, nos résultats viennent corroborer les conclusions auxquelles sont parvenus des chercheurs tel que Koda (1999) pour la reconnaissance de mots en anglais langue seconde.

3.1.1.5 Résultats – Mots discriminants

Dans cette section, nous nous intéresserons aux mots dérivés qui ont contribué à différencier les participants du groupe LBA des participants du groupe LPA. Afin de faciliter la lecture de nos résultats, nous présenterons dans un premier temps les mots dérivés qui ont été les mieux réussis et ceux qui l'ont moins été dans les deux groupes (3.1.1.5.1). Nous pourrions ainsi vérifier s'il existe des similitudes entre les deux groupes et ce, malgré leurs différences statistiquement significatives. Dans un deuxième temps, nous tenterons de dresser un portrait des suffixes discriminant les deux groupes (3.1.1.5.2). Après en avoir dresser la liste pour chaque C-test, nous pourrions non seulement mieux comprendre les suffixes qui posent problème pour des étudiants n'ayant pas une L1 écrite riche en affixes, mais nous pourrions également identifier la nature de ces suffixes (rare ou fréquent? productif ou peu productif?) et vérifier si ceux-ci sont les mêmes pour les deux C-tests.

3.1.1.5.1 Le palmarès des mots dérivés

Le palmarès des mots dérivés du C-test 1

Avant de présenter notre premier tableau qui portera sur les mots dérivés les plus réussis dans le C-test 1, nous tenons à préciser que nous avons décidé de restreindre chaque liste à cinq mots dérivés¹¹. Cette décision a été prise en fonction d'un critère bien précis : le nombre total de mots dérivés que contenait chacun des C-tests. Nous ne voulions pas que le nombre de mots présentés dans la liste des mots les plus réussis et dans celle des mots les moins réussis dépasse cinquante pourcent du nombre total des mots dérivés. Dans le premier C-test, qui est celui qui contenait le moins de mots dérivés parmi les trois, cela représente un maximum de douze à quatorze mots. Étant donné que plusieurs mots dérivés sont présents plus d'une fois dans les C-tests, nous avons choisi de restreindre ce nombre à dix mots, soit cinq mots dérivés pour chaque liste.

Tableau 3.6
Mots dérivés les plus réussis – C-test 1

Groupe LBA	Groupe LPA
Mots dérivés (% de réussite)	
1- difficultés* (96%) (93%) (93%)	1- inscription (83%)
2- adaptation (93%)	2- description (81%)
3- intégration (93%)	3- étudiants (81%)
4- durée (93%)	4- adaptation (78%)
5- inscription (93%)	5- suggestions (75%)

* il y en avait trois dans l'exercice

Nous pouvons tout d'abord constater qu'il y a, dans ces deux listes, deux mots similaires : *adaptation* et *inscription*. Que l'on retrouve ces deux mots dérivés dans chacune des listes n'est pas surprenant puisque le suffixe *-ion* est un suffixe

très fréquent et très productif¹². Nous croyons que c'est également pour cette raison que l'on retrouve d'autres mots avec le suffixe *-ion* dans les deux listes (LBA : *intégration*; LPA : *description, suggestions*).

Deuxième constatation intéressante : nous ne retrouvons que des noms communs dans ces deux listes. Il n'y a ni adjectif ni adverbe. Nous pouvons probablement expliquer cette observation du fait que a) le suffixe *-ion* s'ajoute à une base verbale pour former un nom (et nous avons beaucoup de mots dérivés se terminant avec le suffixe *-ion* dans ces listes) et b) nous avons plus de noms dérivés que d'adjectif ou d'adverbe dérivés dans ce C-test (ration de 5 pour 1). Ce qui est également le cas pour nos deux autres C-tests.

La prochaine observation porte sur la nature des suffixes présents sur cette liste (*-ion*; *-ée*; *-ant*). Quatre-vingt-dix pourcent de ces suffixes ont pour fonction de s'ajouter à une base verbale pour former un nom. Seul le mot dérivé *difficultés* a une base adjectivale auquel on a ajouté un suffixe pour former un nom.

Finalement, nous aimerions porter une attention particulière aux mots suivants : *difficultés* et *durée*.

a) *difficultés*

Comme nous l'avons mentionné plus tôt, il s'agit du seul mot sur la liste qui avait une base adjectivale avant l'ajout du suffixe. Le niveau de difficulté de ce mot, selon nous, était triple : tout d'abord, le suffixe *-ité* n'est pas un suffixe très fréquent; l'ajout du suffixe demande un changement important de l'orthographe de la base; et ensuite, le mot *difficulté* a un compétiteur de la même famille beaucoup plus fréquent, l'adjectif *difficile* (tff : 143,73; tfl : 118,11 comparativement au mot *difficulté* qui a un tff de 17,11 et un tfl de 49,80). Alors

¹¹ Pour obtenir la liste complète des mots dérivés et leur pourcentage de réussite, se référer à l'appendice A.

¹² Voir liste de Catherine-Marie Longtin (2004)

que les participants du groupe LPA ont majoritairement choisi le compétiteur le plus fréquent, soit l'adjectif *difficile*, les participants du groupe LBA n'ont pas éprouvé de problème à compléter la troncation et ont choisi le bon suffixe, avec des taux de réussite de 93% et de 96%. Si le niveau de difficulté de ce mot peut facilement justifier la piètre performance des participants du groupe LPA, qui ont obtenu des moyennes de 53% et de 56%, comment expliquer les bons taux de réussite obtenus par le groupe LBA pour ce même mot ? S'agit-il ici d'une simple question de vocabulaire ? Nos résultats ne nous permettent pas de répondre à cette question.

b) durée

Le mot *durée* est un mot beaucoup moins fréquent que son compétiteur de la même famille, le verbe *durer* (durée : tff : 8,01; tfl : 19,39 ; durer : tff : 68,31; tfl : 101,69). En plus, le suffixe *-ée* est un suffixe plutôt rare bien que productif. Comment alors expliquer qu'il soit dans liste des mots dérivés les plus réussis ? Nous croyons ici que le mot *durée* fait tout simplement partie du vocabulaire connu et fréquent des étudiants universitaires. Cela n'explique toutefois pas pourquoi les étudiants du groupe LPA n'ont pas réussi à choisir le bon suffixe pour ce mot (leur taux de réussite pour ce mot est de 53%). Encore une fois, nos résultats ne nous permettent pas ici d'avancer d'hypothèse.

Passons maintenant au tableau des mots dérivés les moins réussis du C-test 1.

Tableau 3.7
Mots dérivés les moins réussis – C-test 1

Groupe LBA		Groupe LPA	
Mots dérivés (% de réussite)			
1- apprentissage	(41%)	1- exigences	(25%) (28%)*
2- méconnaissance	(52%)	2- absence	(25%)
3- rencontres	(52%)	3- apprentissage	(28%)
4- familiale	(56%)	4- disponibilités	(31%)
5- disponibilités	(56%)	5- familiale	(33%)

* il y en avait deux dans l'exercice

Dans les deux listes, nous retrouvons encore une fois plusieurs mots communs : *apprentissage*, *familiale* et *disponibilités*. De plus, il semblerait que ce soit les suffixes les moins fréquents qui posent le plus de problèmes dans les deux groupes, comme les suffixes :

-*ce* : *méconnaissance*, *absence*, *exigences*

-*ial* : *familiale*

-*ité* : *disponibilités*

Au lieu d'avoir choisi ces suffixes, la plupart des participants de notre recherche ont tout simplement écrit le compétiteur de la même famille le plus fréquent. Par exemple, au lieu d'écrire *apprentissage* (tff : 2,65 ; tfl : 10,47), ils ont écrit *apprendre* (tff : 347,65 ; tfl : 286,69); *disponibles* (tff : 14,46 ; tfl : 12,03) au lieu de *disponibilités* (tff : 0,66 ; tfl : 3,58) et tout cela, faisant fi des tirets disponibles. Ces réponses suggèrent qu'après avoir reconnu la base du mot, les participants ont passé en revue tous les mots qui commençaient par cette base pour finalement choisir le mot qu'ils connaissaient le plus et qui est, la plupart du temps, le mot le plus fréquent. Cela reflète en grande partie les conclusions auxquelles est parvenue l'équipe de Colé et coll. (1989) bien que leur recherche se soit déroulée auprès de locuteurs natifs. Ces auteurs en sont arrivés à la conclusion que les adultes locuteurs natifs du français traitaient les mots morphologiquement

complexes d'une manière très systématique en examinant d'abord la base du mot en question, en passant ensuite en revue tous les mots qui commencent par cette base et en choisissant tout simplement le mot le plus fréquent.

Finalement, après avoir passé en revue les erreurs commises au niveau de ces dix mots, nous avons pu faire le constat que les participants des deux groupes ne faisaient pas le même type d'erreur. Alors que les participants du groupe LPA ont surtout fait des erreurs dans le choix du suffixe (dans 88% des cas), les participants du groupe LBA ont surtout commis des erreurs au niveau de la flexion (dans 60 % des cas). Ce qui nous fait donc dire que la grande différence entre le groupe LPA et le groupe LBA, dans ce test et au niveau des mots dérivés manqués, réside essentiellement dans le fait que les participants du groupe LBA maîtrisent mieux le choix des suffixes que les autres participants. Passons maintenant aux mots dérivés du deuxième C-test afin de savoir si nous allons parvenir aux mêmes conclusions.

Le palmarès des mots dérivés du C-test 2

Tout comme pour les mots dérivés du premier C-test, nous présenterons, sous forme de tableaux, les mots dérivés les mieux réussis (tableau 3.8) et ceux les moins réussis (tableau 3.9) du 2^e C-test. Nous terminerons cette section par un bref retour sur les résultats que nous avons obtenus.

Tableau 3.8
Mots dérivés les plus réussis – C-test 2

Groupe LBA		Groupe LPA	
Mots dérivés (% de réussite)			
1- immigrants	(96%) (96%)*	1- immigrants	(94%) (92%)*
2- linguistique	(96%)	2- connaissance	(94%)
3- population	(96%)	3- population	(89%)
4- catégories	(93%)	4- immigration	(83%) (78%) (78%)**
5- immigration	(93%)	5- international	(69%)

* il y en avait deux dans l'exercice

** il y en avait trois dans l'exercice

Dans un premier temps, nous pouvons observer que plusieurs mots sont communs dans les deux listes : *immigrants*, *population*, *immigration*. Nous croyons que ces mots devaient être tous très connus des participants car s'ils ont été bien réussis, cela n'a pas été le cas pour certains mots de la même famille. Par exemple, dans les deux groupes, les participants ont bien complété les mots *immigrants* et *immigration*, mais n'ont pas réussi à trouver le mot *immigrée* (taux de réussite du mot *immigrée* : 36% LPA et 52% LBA). Nous retrouvons une situation similaire pour le mot *connaissance*. Si les participants du groupe LPA ont complété le mot *connaissance* avec un taux de réussite de 94%, ils n'ont toutefois pas réussi à compléter le mot *méconnaissance* dans le C-test 1 (taux de réussite du mot *méconnaissance* : 47%), qui est pourtant un mot de la même famille. Selon nous, les exemples que nous venons de présenter peuvent être expliqués à l'aide du modèle théorique BIA+ que nous avons présenté dans notre chapitre I. Pour décrire brièvement ce modèle, qui explique le traitement de la morphologie en langue seconde, lorsqu'un sujet participe à une tâche de reconnaissance de mots par exemple, tous les mots connus par cette personne qui commencent par les premières lettres qui lui sont présentées sont activés dans l'espace qu'on appelle le « système d'identification des mots ». Les mots qui auront été le plus activés seront ensuite envoyés au « système tâche - décision » qui, comme son nom

l'indique, aura la tâche de sélectionner le mot correspondant aux besoins et aux attentes du sujet. Ce qu'il y a d'intéressant dans ce modèle, c'est que le système tâche – décision n'a aucune influence sur le système d'identification des mots. Le système tâche -décision peut savoir qu'il a besoin d'un adjectif, mais si le système d'identification ne lui envoie que des noms, il y a à ce moment-là conflit. La personne a alors deux options : choisir quand même un mot parmi ceux proposés par son système d'identification des mots ou tout simplement admettre qu'il ne sait pas.

Pour illustrer ce fonctionnement, prenons un exemple parmi nos mots dérivés les plus réussis. Un sujet doit compléter un mot qui commence par les lettres *immigr*. Le mot recherché est l'adjectif *immigrée*. Le sujet peut connaître les mots *immigration* et *immigrant*, mais ne pas connaître le mot *immigrée*. Son système d'identification des mots envoie donc les deux mots au système tâche – décision qui doit en sélectionner un. Le participant sait peut-être très bien qu'il a besoin d'un adjectif mais comme son système d'identification des mots est indépendant du système tâche – décision et qu'il ne lui a envoyé que ces deux mots, il n'a donc d'autre choix que de choisir l'un de ces mots ou de déclarer qu'il ne connaît pas la réponse¹³. Dans un cas comme dans l'autre, le sujet obtient une mauvaise réponse.

Regardons maintenant les résultats que nous avons obtenus au niveau des mots les moins réussis du deuxième C-test.

¹³ Nous aborderons plus en détails la question des choix des sujets dans la section sur les stratégies (section 3.2.3.2)

Tableau 3.9
Mots dérivés les moins réussis – C-test 2

Groupe LBA		Groupe LPA	
Mots dérivés (% de réussite)			
1- expertise	(11%)	1- incitatives	(0%)
2- incitatives	(19%)	2- migratoires	(0%)
3- diversifie	(22%)	3- expertise	(3%)
4- réfugié	(22%)	4- industrialisés	(6%)
5- industrialisés	(22%)	5- réfugié	(8%)

Malgré la différence au niveau des taux de réussite, nous retrouvons encore une fois les mêmes mots dans les deux listes : *expertise*, *incitatives*, *réfugié*, *industrialisés*. Tous ces mots dérivés ont des suffixes peu fréquents, ce qui semble expliquer pourquoi les étudiants des deux groupes ont eu des problèmes pour compléter ces mots. Ajoutons par ailleurs que dans la majorité des cas, les erreurs commises à l'endroit de ces mots sont essentiellement des erreurs dans le choix du suffixe (dans 60% des cas chez les participants du groupe LPA et dans 80% des cas chez les participants du groupe LBA). Au lieu de choisir le bon suffixe, ces participants ont tout simplement écrit les mots compétiteurs de la même famille les plus fréquents :

Mots à compléter

Incitatives (tff : - ; tfl : -)

Migratoires (tff : - ; tfl : -)

Expertise (tff : 2,65 ; tfl : 0,68)

Industrialisés (tff : 0,06 ; tfl : 0,14)

Mots complétés

incitation (tff : 2,05 ; tfl : 0,54)

migration (tff : 0,66 ; tfl : 2,64)

expert (tff : 22,95 ; tfl : 5,00)

industrielle (tff : 5,42 ; tfl : 7,70)

Lorsque le mot compétiteur était plus court que le mot recherché en terme de lettres, les participants laissaient les tirets vides ou complétaient le mot en le mettant au pluriel ou au féminin (ou au féminin pluriel !) Encore ici, la théorie du modèle BIA+ explique parfaitement les résultats que nous avons obtenus.

Que pouvons-nous conclure des observations que nous avons faites par rapport aux mots dérivés les plus réussis et ceux les moins réussis du premier et du deuxième C-test ? Tout d'abord, dans les deux groupes, il y a beaucoup de similitudes. Ils ont pratiquement raté les mêmes mots et réussi les mêmes mots. Dans le cas des mots réussis, ils étaient soit des mots très fréquents dans le vocabulaire ou soit des mots dont le suffixe était très fréquent (ou les deux). Pour ce qui est des mots manqués, il s'agissait essentiellement de mots plus ou moins fréquents dans le vocabulaire ou de suffixes peu fréquents. Ajoutons également que pour les mots manqués, les deux groupes ont généralement fait le même type d'erreur, soit une erreur dans le choix du suffixe. Au lieu de choisir le bon suffixe, ils ont privilégié le mot compétiteur le plus fréquent.

Si ces observations ont pu, en partie, nous éclairer sur les conditions de réussite des mots dérivés et des connaissances de la dérivation qu'ont les participants de nos deux groupes, deux questions toutefois demeurent. Tout d'abord, si quatre mots sur cinq étaient similaires dans les deux listes, le cinquième mot, lui, était bien souvent parfaitement maîtrisé dans le groupe LBA et complètement raté dans le groupe LPA et ce, même si c'était un mot fréquent. Comment expliquer ce phénomène ? De plus, si nous avons constaté de nombreuses similitudes au niveau des mots les plus réussis et au niveau des mots les plus manqués dans les deux groupes, et que nous avons toujours une différence statistiquement significative au niveau de la performance des mots dérivés, d'où vient cette différence ? Afin de répondre à ces deux questions, nous avons examiné en détails les mots dérivés discriminant les deux groupes. Nous vous présentons les résultats dans la prochaine section.

3.1.1.5.2. Mots dérivés discriminants

Quels mots font la différence au niveau du décodage et de la production de mots dérivés entre notre groupe de participants ayant une L1 écrite riche en affixes et

notre deuxième groupe de participants ayant une L1 écrite pauvre en affixes ? Afin de répondre à cette question, nous avons effectué des analyses ANOVA avec l'ensemble des résultats des deux groupes en ce qui concerne les mots dérivés de nos deux C-tests¹⁴. Voici les résultats :

Tableau 3.10
Mots discriminants – C-tests 1 et 2

C-test 1	C-test 2
Exigences	Industrialisés
Absence	Catégories
Difficultés	Proportion
Pression	Scolarité
Déroulement	Connaissance
Intégration	Linguistique
Inévitables	Concentration
Durée	Multiples
Disponibilités	Diverses
	Incitatives
	Migratoires

Au total, ce sont vingt mots dérivés qui causent la différence significative au niveau de la performance entre le groupe LPA et le groupe LBA. Qu'ont en commun tous ces mots ? Tout d'abord, 60 % de ces mots ont un suffixe peu fréquent et ces mêmes mots ont tous un compétiteur de la même famille plus fréquent qu'eux.

Ex. : **Mots à compléter**

Mots complétés

Difficultés (tff : 17,11 ; tfl : 49,80) = *difficile* (tff : 143,73 ; tfl : 118,11)

Scolarité (tff : 1,27 ; tfl : 0,81) = *scolaire* (tff : 8,13 ; tfl : 14,12)

Industrialisés (tff : 0,06 ; tfl : 0,14) = *industrielle* (tff : 5,42 ; tfl : 7,70)

¹⁴ Pour obtenir la liste complète des mots dérivés avec leur pourcentage de réussite pour chacun des deux groupes, se référer à l'appendice A.

En ce qui concerne la réussite de ces mots, c'est le groupe LBA qui a mieux performé significativement que le groupe LPA. Nous pouvons donc supposer que les participants du groupe LBA connaissaient soit plus de suffixes ou tout simplement plus de mots.

De plus, 60 % de ces mots se sont vu attribuer le mauvais suffixe par les participants du groupe LPA. Ceci vient confirmer notre première observation à savoir que lorsqu'un mot tronqué contient un suffixe peu fréquent, ce suffixe est remplacé par un suffixe plus fréquent et ce, même s'il ne correspond pas au nombre de tirets.

En résumé, ce qui explique la différence significative au niveau de la performance des mots dérivés entre les deux groupes se résume à ceci : des noms communs qui ont un suffixe peu fréquent.

Si nous connaissons maintenant la nature des mots dérivés qui posent problème aux participants du groupe LPA, il serait intéressant de connaître les raisons de cette difficulté. Est-ce à cause de la nature des connaissances morphologiques de ces participants, connaissances qui ne pourraient pas être transférées vers le français langue seconde ? Ou est-ce tout simplement dû à un grave manque de vocabulaire ? Nous tenterons de répondre à cette question avec l'analyse des réponses données par les participants lors des entrevues pour la tâche de rappel stimulé (section 3.2.3). Mais pour l'instant, poursuivons avec les résultats que nous avons obtenus lors du 3^e C-test.

3.1.2 Résultats – C-test 3

Comme nous l'avons expliqué dans l'introduction de la section sur les résultats des C-tests, le 3^e C-test a été avant tout bâti afin de servir d'objet de rappel pour la tâche de rappel stimulé. Néanmoins, nous avons jugé qu'il serait intéressant d'en

présenter les résultats dans cette section dans la perspective où nous pourrions appuyer ou nuancer ce que nous avons observé lors des deux premières épreuves.

Tout comme lors de la présentation des résultats des deux premiers C-tests, nous décrirons d'abord la répartition de nos participants lors de cette troisième épreuve; nous enchaînerons ensuite avec la présentation des résultats et nous terminerons par un bref résumé des observations que nous aurons faites. Mentionnons que nous n'aborderons pas les problèmes encourus lors du codage étant donné que ceux-ci sont de même nature que ceux mentionnés au point 3.1.1.2.

3.1.2.1 Répartition des participants

Contrairement aux deux premières épreuves, la tâche de rappel stimulé n'a pas été administrée en classe. Seuls les participants intéressés à poursuivre l'expérience étaient passés en entrevue en dehors des heures de cours. Par conséquent, au lieu d'une soixantaine de participants, comme ce fut le cas pour les deux premiers C-tests, nous avons recueilli les données de vingt-quatre participants.

La répartition de nos participants sera présentée selon le groupe auquel ils appartiennent, soit le groupe des langues pauvres en affixes (LPA) ou le groupe des langues riches en affixes (LBA).

Tableau 3.11
Répartition des participants selon leur L1 – C-test 3

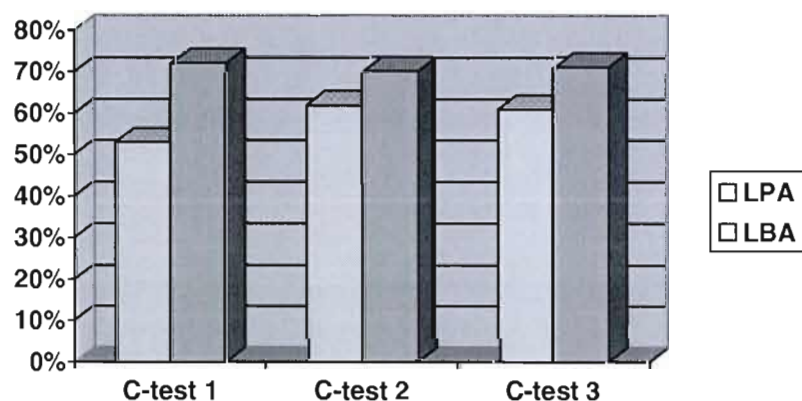
Profil linguistique	<i>Total</i>
LPA	12
LBA	12
<i>Total</i>	<i>24</i>

Nous pouvons d'abord constater que nous avons interviewé autant de participants ayant une L1 écrite pauvre en affixes que de participants ayant une L1 riche en affixes. Ce sont également les étudiants les plus forts de chacun des deux groupes qui se sont présentés à cette dernière épreuve (ces étudiants, à l'exception près de deux, ont obtenu les meilleurs résultats au niveau de la performance globale lors des deux premiers C-tests dans leur groupe respectif).

3.1.2.2 Résultats – Réussite globale

Afin que les résultats que nous avons obtenus lors du 3^e C-test soient réellement significatifs pour notre recherche, nous avons pensé qu'il serait intéressant de les présenter en parallèle avec ceux obtenus lors des deux premières épreuves. Pour ce faire, nous avons refait des analyses des C-tests 1 et 2 avec pour seules données celles de nos vingt-quatre participants.

Figure 3.1
Réussite globale – C-test 3 – LPA \ LBA



Au niveau des caractéristiques morphologiques du système d'écriture de la L1 des participants, on constate que les résultats lors du 3^e C-test ressemblent beaucoup aux résultats des deux premières épreuves, avec une légère supériorité des participants ayant une L1 écrite riche en affixes sur ceux ayant une L1 pauvre en

affixes. Toutefois, contrairement aux deux premières épreuves, cette différence de 10% n'est pas statistiquement significative (df 1; F 3,404; Sig. 0,079). Nous pensons que cela est dû, en partie, au fait que ce sont les étudiants les plus forts des deux groupes qui ont choisi volontairement de participer au 3^e C-test. La différence aurait probablement été significative, tout comme elle l'a été lors des deux premières épreuves, si nous avions administré ce troisième C-test à tous nos participants.

3.1.2.3 Résultats – Réussite des mots dérivés

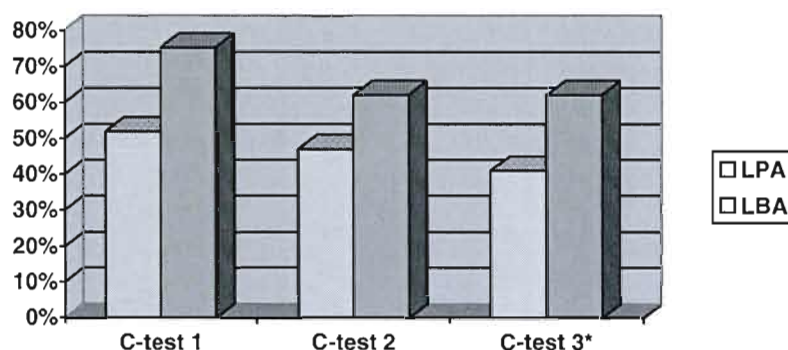
Comme nous l'avons mentionné dans l'introduction de la section 3.1.2, les résultats que nous présenterons ici nous permettront, en quelque sorte, d'appuyer ou de nuancer ce que nous avons observé lors de nos deux premières épreuves. Afin de faciliter ce type d'observation, et comme nous l'avons fait pour la réussite globale, nous présenterons les taux de réussite des mots dérivés du 3^e C-test en parallèle avec ceux obtenus lors des deux premiers C-tests. Les résultats obtenus seront présentés selon le groupe (LPA \ LBA).

Terminons cette brève introduction en mentionnant que le 3^e C-test compte vingt-sept mots dérivés sélectionnés en vue de représenter la plus grande variété de suffixes possibles¹⁵.

¹⁵ Par « variété », nous entendons ici une sélection de suffixes qui s'est faite à deux niveaux : au niveau de leur productivité et de leur fréquence.

Figure 3.2

Réussite des mots dérivés – C-test 3 – LPA \ LBA

* significatif $\leq 0,05$

La moyenne obtenue par le groupe LBA lors de la troisième épreuve vient confirmer la supériorité de ce groupe en ce qui a trait à la réussite des mots dérivés. Non seulement ont-ils obtenu de meilleurs résultats que leurs camarades, mais en plus, à toutes les épreuves, la différence entre leurs résultats et ceux du groupe LPA est statistiquement significative (pour le 3^e C-test, la différence entre les moyennes est de 21% et est statistiquement significative selon des analyses ANOVA (df 1; F 12,735; Sig. 0,002)).

Ceci vient donc appuyer notre première hypothèse de recherche qui stipulait que des apprenants ayant une L1 écrite riche en affixes auraient de meilleurs résultats que des apprenants ayant une L1 écrite pauvre en affixes dans un exercice de décodage et de production de mots dérivés en français langue seconde. Cette conclusion vient corroborer les résultats obtenus par de nombreux chercheurs pour l'anglais langue seconde (se référer au chapitre I).

Si les résultats que nous avons obtenus lors du 3^e C-test viennent appuyer nos observations initiales, dans la même perspective, il serait intéressant d'analyser les mots discriminants du 3^e C-test. C'est ce que nous nous proposons de faire dans la section suivante.

3.1.2.4 Réussite – Mots discriminants

Dans cette section, nous présenterons dans un premier temps les mots dérivés les plus réussis et ceux les plus manqués du C-test 3 (3.1.2.4.1); et dans un deuxième temps, les mots discriminant les deux groupes (3.1.2.4.2). Bien entendu, les listes de mots que nous obtiendrons ne seront pas identiques à celles des deux premiers C-tests, étant donné que les textes des épreuves étaient différents. Dans ce cas-ci, nous chercherons à vérifier si nous allons observer les mêmes constantes au niveau des suffixes. Est-ce que les mots généralement les plus réussis sont ceux dont le suffixe est très fréquent ? Est-ce que les mots les moins réussis sont ceux dont le suffixe est moins fréquent ? Est-ce que les mots dérivés discriminant les deux groupes sont des noms communs dont le suffixe est peu fréquent ? La présente section nous permettra de répondre à ces questions.

3.1.2.4.1 Le palmarès des mots dérivés du 3^e C-test

Nous présenterons d'abord les cinq mots dérivés les plus réussis (tableau 3.12) et ensuite les cinq mots dérivés les plus manqués (tableau 3.13). Pour une justification du nombre de mots présentés par catégorie, se référer à la section 3.1.1.5.1.1.

Tableau 3.12
Mots dérivés les plus réussis – C-test 3

Groupe LBA		Groupe LPA	
Mots dérivés (% de réussite)			
1- capacité	(100%)	1- scolaire	(92%)
2- concentration	(100%) (100%)	2- capacité	(92%)
3- planification	(100%)	3- simplement	(83%)
4- perfection	(92%)	4- durée	(83%)
5- pressions	(92%)	5- planification	(83%)

* il y en avait deux dans le texte

Ce tableau nous permet de constater qu'il existe des similitudes entre les deux groupes. Tout d'abord, nous retrouvons des mots dérivés communs aux deux listes : *capacité* et *planification*. Ensuite, 60% de ces mots dérivés ont des suffixes fréquents (-ion; -ité). Pour ce qui est des mots dont les suffixes sont d'une fréquence moyenne (*capacité*; *durée* et *scolaire*), nous pouvons observer que ce sont des mots qui, sans avoir un suffixe fréquent, sont eux-mêmes fréquents dans le vocabulaire scolaire, et par conséquent, dans le vocabulaire de nos participants. Ceci expliquerait, en partie, pourquoi ces mots se retrouvent dans la liste des mots dérivés les plus réussis.

Nous aimerions terminer ces premières observations en traitant d'un mot dérivé spécifique : le mot *durée*. Ce mot était également présent dans le premier C-test. Alors qu'il avait été très bien réussi par les participants du groupe LBA (93%), les participants du groupe LPA avaient obtenu une moyenne de 53% pour ce même mot. Or, dans le troisième C-test, le mot *durée* se retrouve dans la liste des mots dérivés les plus réussis par le groupe LPA (à noter qu'il a également été très bien réussi par le groupe LBA avec un taux de réussite de 92%). Que s'est-il passé ? Comme nous l'avons expliqué dans notre section décrivant la répartition des participants du 3^e C-test, les étudiants qui se sont portés volontaires sont ceux qui ont le mieux performé lors des deux premiers C-tests et ce, dans leur groupe respectif. Il n'est donc pas surprenant que les participants du groupe LPA aient obtenu de meilleurs résultats lors du 3^e C-test, d'où le taux de réussite élevé pour le mot *durée*.

Tableau 3.13
Mots dérivés les plus manqués – C-test 3

Groupe LBA		Groupe LPA	
Mots dérivés (% de réussite)			
1- échéancier	(0%)	1- échéancier	(0%)
2- contingentement	(0%)	2- contingentement	(0%)
3- stresseurs	(0,08%)	3- stresseurs	(0%)
4- démotivant	(0,17%)	4- démotivant	(0%)
5- procrastination	(0,17%)	5- procrastination	(0%)

Comme nous pouvons le constater, les mots des deux listes ci-dessus sont parfaitement identiques. Ce sont donc les mêmes mots dérivés qui ont occasionné des problèmes aux deux groupes d'étudiants. Puisque que tous ces mots ont des suffixes fréquents (-ion; -ment; -eur; -ant), pourquoi les étudiants ne les ont pas réussis ? Afin de répondre à cette question, nous avons examiné les raisons pour lesquelles ces mots avaient été ratés. Ce que nous avons pu observer, c'est que si les participants se sont trompés de suffixes dans certains cas, ils ont surtout choisi tout simplement de ne pas compléter ces mots. Ceci nous porte donc à croire que les participants ne connaissaient tout simplement pas ces mots.

En résumé, nos dernières observations sur les mots dérivés viennent encore une fois confirmer ce que nous avons observé lors de l'analyse des deux premiers C-tests à savoir que les deux groupes d'apprenants LPA et LBA ont réussi et ont raté en général les mêmes mots. Les mots les plus réussis sont a) des mots fréquents dans le vocabulaire de nos participants ou b) des mots dont le suffixe est très fréquent. Pour ce qui est des mots les plus ratés, même s'ils avaient un suffixe fréquent, ils étaient peu connus des participants, la preuve étant que les participants ont majoritairement choisi de ne pas les compléter.

Finalement, mentionnons que nous n'avons constaté qu'une seule différence entre les analyses effectuées lors du 3^e C-test et celles effectuées lors des deux

premières épreuves. Si, dans le 1^{er} et le 2^e C-test, l'erreur la plus commune a été de choisir le mauvais suffixe, dans le 3^e C-test, les participants ont choisi de ne pas compléter les mots. L'ont-ils fait par manque de connaissances lexicales ? C'est ce que nous verrons lors de l'analyse des stratégies identifiées lors de la tâche de rappel stimulé (3.2.3).

S'il existe plusieurs similitudes entre les deux groupes pour ce qui est des mots les plus réussis et si les mots manqués sont exactement les mêmes, quels sont donc les mots qui justifient la différence statistiquement significative entre ces deux groupes ? Nous avons découvert lors de l'analyse du premier et du deuxième C-test que ce sont surtout les noms communs ayant un suffixe peu fréquent qui discriminaient les deux groupes. Pouvons-nous confirmer la même observation à l'aide des résultats du 3^e C-test ? C'est ce que nous tenterons de faire dans la prochaine section.

3.1.2.4.2 Mots dérivés discriminants

Quels mots dérivés font la différence entre le groupe LPA et le groupe LBA au niveau du taux de réussite ? Afin de répondre à cette question, nous avons effectué des analyses ANOVA dans le but de vérifier si des mots se démarquaient statistiquement des autres¹⁶. Voici les résultats que nous avons obtenus :

¹⁶ Pour obtenir la liste complète des mots dérivés avec leur pourcentage de réussite pour chacun des deux groupes, se référer à l'appendice A.

Tableau 3.14
Mots discriminants – C-test 3

C-test 3
Parentales
Répartition
Efficacement
Concentration

Au total, 14% des mots dérivés du troisième C-test discrimine les deux groupes. Contrairement aux mots discriminants des deux premières épreuves, la majorité de ces mots ont un suffixe fréquent. Ceci n'est toutefois pas suffisant pour que les participants du groupe LPA les aient réussis. Dans le cas du troisième C-test, la majorité des erreurs commises (50%) n'a pas été au niveau du choix du suffixe, mais bien parce que le mot n'a tout simplement pas été complété. L'autre erreur est de l'ordre de la dérivation (dans le mot *parentales*).

En résumé, les mots dérivés discriminant les deux groupes seraient des mots peu connus (ou tout simplement pas connus) des participants du groupe LPA. Si les résultats du troisième C-test semble nous orienter vers ce constat, l'analyse des réponses aux entrevues de la tâche de rappel stimulé, que nous vous présentons dans la prochaine section, devrait être en mesure de nous éclairer davantage sur la question.

3.2 Résultats – Tâche de rappel stimulé

Rappelons d'abord brièvement la nature de la tâche de rappel stimulé et les conditions dans lesquelles elle s'est déroulée.

La tâche de rappel stimulé est un exercice d'introspection qui permet d'explorer les pensées ou stratégies d'un apprenant au moment où il accomplissait une

activité. Le souvenir des processus cognitifs est stimulé à l'aide d'un rappel, visuel ou oral, et de questions neutres auxquelles l'apprenant doit répondre à haute voix. Les réponses sont enregistrées.

Dans le cadre de notre recherche, l'objet qui nous a servi de rappel est le troisième C-test. Lors d'entrevues individuelles, qui se sont déroulées en dehors des heures de cours, les participants ont d'abord été invités à compléter un C-test. Une fois l'exercice terminé, nous avons procédé à la tâche de rappel stimulé en leur posant des questions sur les mots qu'ils venaient de compléter. Ces questions visaient essentiellement à connaître les stratégies qu'ils avaient employées pour compléter les mots dérivés tronqués. Pour obtenir davantage d'informations sur le protocole de recherche de la tâche de rappel stimulé, veuillez consulter l'appendice B1.

Avant de passer à la présentation de nos résultats (point 3.2.3), nous souhaitons d'abord décrire la répartition de nos participants (3.2.1) ainsi que les problèmes que nous avons rencontrés lors de l'administration de la tâche de rappel stimulé (3.2.2).

3.2.1 Répartition des participants

Au total, vingt-quatre personnes de notre échantillon initial se sont portés volontaires pour participer à la tâche de rappel stimulé. Pour une description détaillée de la répartition de nos participants, se référer à la section décrivant la répartition des participants du 3^e C-test (3.1.2.1.).

3.2.2 Problèmes encourus

Les problèmes que nous avons rencontrés lors de la tâche de rappel stimulé concernent surtout le déroulement de l'épreuve en question. Pour ce qui est du codage des réponses des participants, nous n'avons eu aucune difficulté et ce,

grâce à la finesse de notre grille de codage ainsi qu'aux enregistrements des réponses, ce qui nous a permis de réécouter les entrevues en cas de doute.

Globalement, nous avons rencontré trois types de problèmes lors de l'administration de la tâche de rappel stimulé. Ces problèmes, bien que mineurs, ont demandé des ajustements de la part de l'interviewer à chaque fois qu'ils se présentaient. Nous vous les décrivons ici sommairement avec les solutions que nous avons trouvées pour y remédier. Mentionnons qu'ils ne sont pas présentés selon un ordre d'importance.

a) Présence du magnétophone

L'un des premiers problèmes auxquels l'interviewer a dû faire face est la gêne que provoquait chez le sujet la présence du magnétophone. Cette présence a intimidé plusieurs participants, surtout au début de l'entrevue. Les réactions observées étaient alors diverses : certains participants ne savaient plus quoi répondre; d'autres ne savaient plus soudainement parler français; ils regardaient sans cesse le magnétophone; ils figeaient sans être capables de dire quoique ce soit ; etc. Pour pallier à ce problème, une fois le consentement verbal enregistré, l'interviewer cachait le magnétophone derrière une plante qui se trouvait sur le bureau. Ainsi, le participant savait que ses réponses étaient enregistrées, car il avait vu le magnétophone, mais il n'avait plus à être gêné par la présence de l'objet durant l'entrevue.

b) Faible niveau du français oral des participants

Bien que l'on ait essayé d'éviter ce problème en choisissant des participants du niveau avancé de français écrit, il s'avère que plusieurs d'entre eux étaient très faibles au niveau du français oral. Non pas que nous ayons eu de la difficulté à comprendre leurs réponses, mais plutôt que nous nous sommes interrogés sur leur capacité à utiliser un métalangage. Par exemple, lorsqu'un participant nous répondait qu'il ne savait pas comment il avait trouvé ce mot, est-ce parce qu'il ne se rappelait pas vraiment comment il avait trouvé ce mot ou est-ce plutôt parce

qu'il ne connaissait tout simplement pas les mots pour dire comment il s'y était pris ? Étant donné que nous n'avions aucun moyen de savoir si, réellement, les participants n'avaient pas les mots pour décrire les stratégies qu'ils avaient utilisées, et puisque nous ne pouvions pas leur donner d'exemples sans biaiser nos données, nous avons tout simplement coché « ne sait pas » comme stratégie identifiée par le participant. Nous aborderons plus en détails cette stratégie du « ne sait pas » quand nous traiterons des résultats (3.2.3.1).

c) Désir de plaire à l'interviewer

Gass et Mackey, dans leur livre sur la tâche de rappel stimulé (2000), avaient parlé du « désir de plaire à l'interviewer » comme d'un piège dans lequel l'interviewer devait absolument éviter de tomber. Afin d'empêcher cette situation de se produire, nous avons bâti un protocole très rigoureux permettant ainsi à l'interviewer de limiter au maximum ses interventions et ainsi ne pas favoriser ce comportement par des encouragements excessifs. Malgré cela, nous avons observé ce comportement, ce désir de plaire à l'interviewer, chez certains de nos participants. Ce comportement se traduisait surtout par de longues explications qui visaient essentiellement à obtenir l'approbation de l'interviewer. Ces explications pouvaient être en lien avec le sujet ou diverger pour complimenter l'interviewer sur sa recherche, la qualité de son français, etc. Si nous n'avons pas pu prévenir ce comportement, nous avons pu néanmoins limiter son effet à une ou deux questions et ce, grâce à la rigueur de notre protocole. Le candidat, voyant que ses commentaires n'avaient aucun effet sur l'interviewer et que celui-ci ne l'aiderait pas davantage, changeait tout simplement d'approche. La rigueur du protocole de recherche a pris ici toute son importance.

3.2.3 Résultats - Tâche de rappel stimulé

Puisque seule la variable « caractéristiques morphologiques du système d'écriture de la langue première » a eu un impact non négligeable sur les résultats des

participants aux C-tests, nous avons cru qu'il serait intéressant de présenter les données recueillies lors de la tâche de rappel stimulé dans cette même perspective, c'est-à-dire en comparant les réponses des participants du groupe LPA et celles du groupe LBA. Nous présenterons ces résultats sous forme de figures et de tableaux. Chaque tableau et/ou figure sera suivi de commentaires. Nous présenterons d'abord les catégories de stratégies les plus mentionnées par les deux groupes de participants (3.2.3.1). Nous pourrions ainsi vérifier s'il existe des similitudes et des différences entre les deux groupes. Ensuite, nous décrirons les stratégies discriminant les deux groupes (3.2.3.2) et nous examinerons la correspondance entre les stratégies identifiées par les participants et les réponses qu'ils ont écrites (3.2.3.3). Nous terminerons par un bref résumé des observations que nous avons faites.

Avant de passer aux premiers résultats, nous aimerions mentionner que nous n'avons effectué aucune analyse statistique avec les données que nous avons recueillies lors de la tâche de rappel stimulé. Nous avons jugé que le nombre de participants n'était pas assez élevé pour pouvoir généraliser nos observations. Néanmoins, nous croyons que les résultats que nous présenterons nous permettront d'apporter une certaine lumière sur les stratégies utilisées par des adultes d'horizons linguistiques différents lorsqu'ils décodent et produisent des mots dérivés en français langue seconde.

3.2.3.1 Résultats – Catégories de stratégies les plus nommées

Pour faciliter le codage et la présentation des stratégies les plus nommées par les participants, nous les avons regroupées en quatre catégories (voir tableau 2.4, p.50). Nous jugeons important de faire un retour ici sur la catégorie « ne sait pas » étant donné le nombre assez important de réponses comme celle-là recueillies lors des entrevues. Quand un participant nous répondait qu'il ne savait pas comment il avait trouvé ce mot, il était difficile pour nous de savoir s'il ne

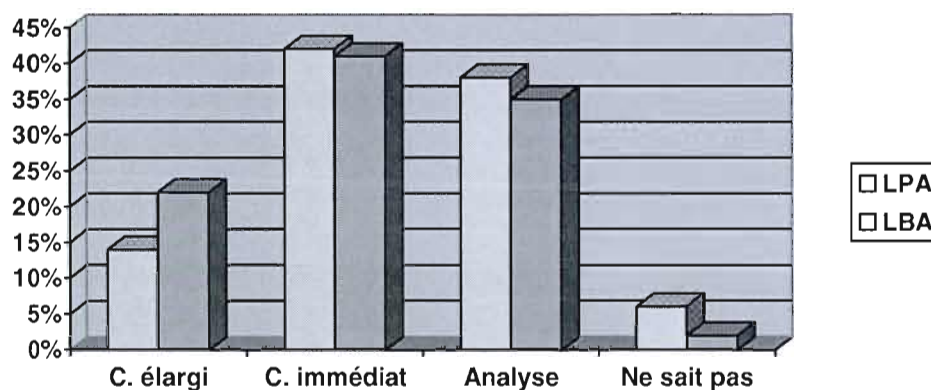
savait réellement pas comment il avait trouvé la réponse ou s'il manquait tout simplement de mots pour le dire. À ce sujet, Feldman et Stemmer (1987) proposent une autre raison pour expliquer cette incapacité chez le participant à identifier la stratégie qu'il a utilisée :

« Another possibility is that problem-solving simply did not take place in a verbal form and consequently, there was nothing to be verbalized. » (p.263)

Que le sujet ait tout simplement complété les mots sans vraiment y penser est une explication fort probable, mais que nous ne pouvons pas vérifier. Ceci étant dit, il sera intéressant de voir de quelle manière cette stratégie se situe par rapport aux autres stratégies et si un groupe l'a davantage utilisée.

Quelles sont les catégories de stratégies les plus nommées ? Voici une figure présentant nos compilations.

Figure 3.3
Catégories de stratégies les plus identifiées et celles les moins nommées



Première constatation : le contexte immédiat est la catégorie de stratégies la plus identifiée lors de la tâche de rappel. Ce résultat n'est pas surprenant et il avait d'ailleurs été constaté par Koda (1999) dans une recherche sur la reconnaissance de mots en anglais langue seconde. En effet, ce dernier, dans une recherche

menée auprès de collégiens coréens et japonais, avait constaté que s'il existait des différences au niveau des stratégies déployées par les sujets pour accomplir une tâche de reconnaissance de mots, il existait également une similitude importante qui était que tous ses sujets s'étaient fiés aux premières lettres données des mots ainsi qu'aux mots environnants pour compléter le mot tronqué. Nous arrivons à la même conclusion avec nos résultats. Peu importe les caractéristiques morphologiques du système d'écriture de leur langue maternelle, tous nos participants se sont d'abord fiés au contexte immédiat pour compléter les mots tronqués.

La deuxième catégorie de stratégies la plus nommée est l'analyse grammaticale et lexicale. Cette stratégie a été identifiée sensiblement dans les mêmes proportions dans les deux groupes. Encore une fois, nous ne sommes pas surpris de ce résultat étant donné que nous avons mené notre recherche dans un milieu scolaire. De plus, nos sujets étaient en train d'étudier le français écrit et par conséquent, les règles de grammaire régissant l'écrit. Il n'est donc pas surprenant que l'analyse grammaticale et lexicale occupe la deuxième place. Peut-être en aurait-il été autrement si nous avions mené notre recherche auprès d'adultes travaillant dans une usine par exemple.

Finalement, si nous mettons de côté la catégorie du « ne sait pas », nous avons en dernière place la catégorie du contexte élargi. Comparativement aux autres catégories, le contexte élargi a été très peu utilisé par nos participants. Ils ne se sont pas vraiment fiés au sens général du texte, au sens du paragraphe ou au sens de la phrase. Ce résultat nous a surpris étant donné que le C-test est avant tout un texte à trous et qu'il fallait comprendre à tout le moins le sens global du texte pour avoir une idée du sujet et être ainsi en mesure de compléter les mots. Ceci expliquerait, en quelque sorte, les réponses tout à fait hors sujet que certains participants ont écrit dans leurs C-tests (ex. : *facture* au lieu de *facteurs*). Ils n'ont cherché qu'à compléter les mots et l'ont fait à l'aide de mots qu'ils connaissaient dans leur vocabulaire. Cependant, comme l'ont fait remarquer

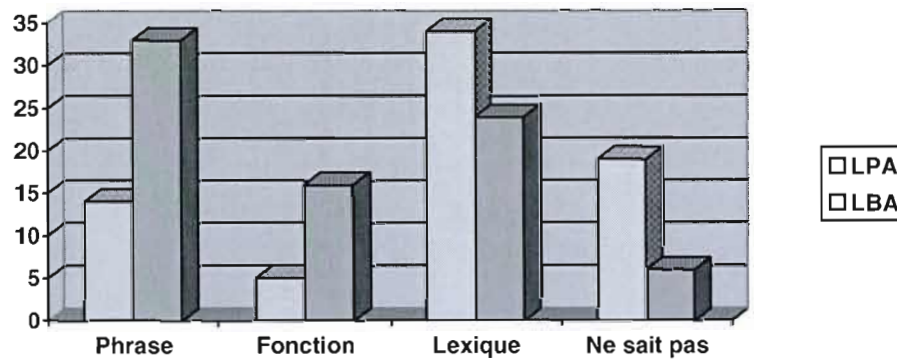
Feldman et Stemmer (1987), il est très difficile pour des apprenants de langue seconde de faire attention à la fois au mot à compléter et au sens de la phrase ou du texte et ce, à cause de la surcharge cognitive. Selon elles, ce n'est que si on acquière une certaine compétence dans la langue seconde que l'on peut se détacher du contexte immédiat pour aller chercher des indices dans le contexte élargi, c'est-à-dire dans le sens du texte ou du paragraphe, et dans les redondances naturelles du texte. Si on se fie aux données que nous avons recueillies, malgré leur niveau avancé en français écrit, la plupart de nos participants ne semblent pas avoir encore acquis cette compétence.

Pour faire un court résumé, les deux groupes ont sensiblement identifié les mêmes stratégies pour compléter les mêmes mots et ce, dans des proportions similaires. Si les participants du groupe LPA ont dit s'être fiés davantage (mais dans une proportion assez faible) sur le contexte immédiat et l'analyse grammaticale et lexicale pour compléter les mots, les participants du groupe LBA, eux, ont mentionné plus souvent le contexte élargi. Est-ce parce qu'ils auraient une meilleure compétence générale en français que leurs camarades de classe (hypothèse de Feldman et Stemmer) ? Cela viendrait-il expliquer la différence statistiquement significative au niveau des résultats obtenus lors du décodage et de la production de mots dérivés ? Dans la prochaine section, nous allons examiner de plus près les occurrences des stratégies mentionnées afin d'identifier celles qui différencient le plus les deux groupes.

3.2.3.2 Stratégies discriminant les deux groupes

Nous tenons d'abord à rappeler que nous n'avons fait aucune analyse statistique pour ce qui est des données que nous avons recueillies lors de la tâche de rappel stimulé (voir justification point 3.2.3). Les stratégies que nous allons présenter comme discriminantes ont été choisies parce qu'elles représentaient la plus grande différence, au niveau du nombre d'occurrences, entre les deux groupes.

Figure 3.4
Stratégies discriminantes



Voici les quatre stratégies pour lesquelles on constate la plus grande différence entre les deux groupes au niveau des occurrences : le sens de la phrase, la fonction du mot à compléter, la recherche dans le lexique mental et le « ne sait pas ». Dans cette section, nous allons aborder en détails chacune de ces stratégies.

a) Le sens de la phrase

Dans la section 3.2.3., nous avons pu observer que c'est la catégorie du contexte élargi qui avait été la moins mentionnée lors des entrevues et ce, dans les deux groupes. Nous avons également constaté que c'était dans cette même catégorie que l'on retrouvait la plus grande différence au niveau des occurrences entre les deux groupes. Mais cette catégorie regroupant quatre stratégies spécifiques, il nous était difficile de savoir s'il y avait une stratégie en particulier qui discriminait véritablement les deux groupes. La figure 3.4 vient apporter un nouvel éclairage à la question. En effet, c'est la stratégie en lien avec le sens de la phrase qui vient différencier les deux groupes, les participants du groupe LBA l'ayant mentionnée plus souvent que les participants du groupe LPA.

Si les participants du groupe LBA y ont eu plus souvent recours, est-ce parce qu'ils ont acquis un certain niveau de compétence en français qui leur permet de se distancier du contexte immédiat du mot et de chercher des indices ailleurs dans

le texte, comme le stipulent Feldman et Stemmer ? Si tel est le cas, quelle compétence ont-ils acquise qui ne serait pas encore maîtrisée par les participants du groupe LPA ? Les deux groupes de participants, rappelons-le, ont sensiblement le même niveau en français. Qu'entendent Feldman et Stemmer par « compétence » ? Nous trouverons peut-être la réponse à ces questions lorsque nous aborderons les autres stratégies discriminantes.

b) La fonction du mot à compléter

La stratégie de la fonction du mot à compléter fait partie de la catégorie de stratégies appelée « analyse grammaticale et lexicale ». Le participant qui utilise cette stratégie démontre, en quelque sorte, qu'il est capable de se poser des questions par rapport au texte, mais surtout par rapport à la phrase, des questions telles que : ai-je besoin d'un sujet ou d'un verbe ? Il démontre également qu'il est capable d'analyser la phrase afin d'en reconnaître la structure de base et d'identifier le ou les éléments manquant(s). Nous ne sommes donc pas étonnés si ce sont les participants du groupe LBA qui ont mentionné le plus souvent cette stratégie étant donné que ce sont eux qui ont eu le plus souvent recours au sens de la phrase. Selon nous, cette stratégie de la fonction du mot va de pair avec la stratégie du sens de la phrase.

Si cette observation ne nous étonne pas, nous sommes toutefois surpris par la différence assez importante du nombre d'occurrences entre les deux groupes (LPA : 5; LBA : 16), surtout que nous avons initialement observé un nombre similaire d'occurrences pour la catégorie de l'analyse grammaticale et lexicale (voir figure 3.3). Si le nombre de références à la catégorie de l'analyse grammaticale et lexicale est similaire pour les deux groupes et que la stratégie de la fonction du mot, qui fait partie de cette catégorie de stratégies, a surtout été nommée par le groupe LBA, il doit y avoir une deuxième stratégie dans cette même catégorie qui a été plus souvent mentionnée par le groupe LPA afin de compenser. Il semble que ce soit effectivement le cas, comme nous pourrions le constater avec notre troisième stratégie discriminante.

c) Recherche dans le lexique mental

La stratégie de la recherche de mots dans son lexique mental fait partie de la catégorie de l'analyse grammaticale et lexicale. Comme nous le démontre notre graphique, ce sont surtout les participants du groupe LPA qui y ont fait le plus souvent référence. Elle vient donc faire contrepoids à la stratégie de la fonction du mot, dans la même catégorie, qui a été surtout mentionnée par les participants du groupe LBA.

À quoi correspond exactement la stratégie de la recherche de mots dans le lexique mental ? C'est une stratégie qui sous-entend que le sujet recherche le mot à compléter parmi l'ensemble des mots qu'il connaît. La recherche dans le lexique mental se fait, bien sûr, à partir d'indices textuels dont, avant tout autre indice, les premières lettres données du mot. Voici un exemple de réponse donnée par un participant qui illustre bien cette stratégie.

Ex. : Je vois les lettres « *diffi* » et je pense à tous les mots que je connais qui commencent par ces lettres. (Participant 111)

Que ce soit surtout les participants du groupe LPA qui aient fait mention de cette stratégie révèle, entre autres, la manière dont ils ont procédé pour compléter les mots du C-test. Avant même de penser à la fonction du mot ou avant même de comprendre le sens de la phrase, ces participants ont passé en revue tous les mots de vocabulaire qu'ils connaissaient ayant les mêmes premières lettres que celles données dans le C-test. Cette manière de procéder n'est pas sans rappeler celle décrite par Colé et coll. (1989) lors d'une recherche menée auprès de locuteurs natifs du français. Les auteurs avaient constaté que tous les adultes traitaient les mots morphologiquement complexes de la même manière, soit en identifiant d'abord la base et en passant en revue, par la suite, tous les mots qu'ils connaissaient ayant cette même base. Une question toutefois demeure : pourquoi les participants du groupe LPA n'ont-ils pas poursuivi leur questionnement une fois la recherche dans leur lexique mental effectuée ? Pourquoi ne se sont-ils pas ensuite interrogés sur la fonction du mot à compléter ou même sur le sens de la

phrase dans lequel est inscrit le mot à compléter ? Nous croyons que la réponse se trouve dans la richesse même du lexique mental des participants et non pas dans leur capacité à analyser. Pour expliquer notre hypothèse, faisons un bref retour en arrière et examinons encore une fois notre modèle théorique, le modèle BIA+.

Selon le modèle BIA+, le système d'identification des mots est indépendant du système tâche - décision. Lorsqu'une personne fait une recherche dans son lexique mental, elle active, selon les indices qu'elle aura recueillis, un ensemble de mots dans les langues qu'elle maîtrise. Cet ensemble de mots est alors retransmis au système tâche - décision qui aura comme tâche de sélectionner le mot approprié selon les objectifs de la personne en question. Or, il peut arriver que le système d'identification des mots ne transmet qu'un seul mot au système tâche - décision ou, pire scénario, qu'il ne trouve tout simplement pas de mots correspondant aux indices donnés. La personne se retrouve donc avec deux options : soit elle choisit le mot que le système d'identification des mots lui a transmis et ce, même si ce mot ne correspond pas à ses attentes, soit elle décide de ne rien choisir du tout.

Si nous nous fions à ce modèle et si les participants du groupe LPA ont effectivement un lexique mental plus pauvre que les participants du groupe LBA, il devient alors facile d'identifier la raison pour laquelle ils n'ont pas poursuivi leur questionnement : d'autres indices ne les auraient tout simplement pas aidés puisqu'ils ne connaissaient pas d'autres mots. La richesse lexicale pourrait donc être l'élément clé différenciant nos deux groupes de participants (rappelons que les résultats du 3^e C-test nous orientaient dans cette direction).

Nous pensons également qu'il est fort probable que la « compétence » dont parlaient Feldman et Stemmer se définit, en partie, par la richesse des connaissances lexicales. Si nous acceptons cette hypothèse, les participants du groupe LBA auraient de meilleures connaissances lexicales que les participants du groupe LPA, ce qui a été en partie démontré par notre analyse des mots

discriminants du 3^e C-test. Par conséquent, puisque leur système d'identification des mots leur aurait retransmis une plus grande quantité de mots, ils auraient eu besoin de trouver davantage d'indices pour sélectionner le mot approprié. Ceci expliquerait pourquoi ils ont eu plus souvent recours aux stratégies du sens de la phrase et à la fonction du mot. Bien que fort probable, ceci demeure une hypothèse. Nous pensons toutefois être en mesure de la vérifier et ce, de deux façons :

- 1- Nous allons vérifier la correspondance entre les stratégies mentionnées lors des entrevues et les réponses écrites dans le 3^e C-test. Pour ce faire, nous allons surtout examiner la correspondance entre la stratégie du lexique mentionnée par les participants du groupe LPA et les réponses que ces mêmes participants ont écrites. Il est possible que les participants du groupe LPA aient choisi la stratégie de la recherche dans le lexique mental et aient obtenu de bonnes réponses. Mais il est également possible, et c'est ce qui viendrait confirmer notre hypothèse, que ces participants aient identifié la stratégie du lexique mais qu'ils aient écrit la mauvaise réponse. Nous trouverions là une première évidence que le système d'identification des mots est indépendant du système tâche - décision et que, bien que le sujet connaisse et utilise de bonnes stratégies, il écrit la mauvaise réponse par manque de vocabulaire. Ce constat viendrait également suggérer que les participants du groupe LPA ont des connaissances lexicales plus pauvres que celles de leurs camarades de classe. L'analyse des correspondances entre les stratégies mentionnées lors des entrevues et les réponses écrites dans le C-test seront présentées dans la section 3.2.3.3.
- 2- Une autre manière de vérifier cette hypothèse est d'examiner plus en détails la stratégie du « ne sait pas ». Incidemment, cette stratégie fait partie des stratégies discriminant les deux groupes. Elle est d'ailleurs le sujet de notre prochain paragraphe.

d) La stratégie du « ne sait pas »

Tel que discuté au point 3.2.3.1, nous avons dès le départ des réserves quant à la pertinence de prendre en considération cette stratégie (qui n'en est pas une en fait !) lors de la compilation de nos données. Néanmoins, nous avons tout de même décidé de la considérer dans nos calculs et il se trouve que celle-ci fait partie des stratégies qui discriminent nos deux groupes de participants.

La figure 3.4 illustre très clairement que les participants du groupe LPA ont mentionné plus souvent cette stratégie que les participants du groupe LBA (une différence de treize mentions). Afin d'en apprendre davantage, nous avons voulu vérifier si ceux qui avaient mentionné cette stratégie avaient écrit de bonnes réponses ou s'ils avaient choisi de mauvaises réponses.

Chez les participants du groupe LBA, dans 67 % des cas, les mots écrits étaient les bons alors que le taux de réussite oscille autour du 42% chez les participants du groupe LPA, ce qui est nettement plus faible. Si les participants du groupe LPA ont écrit de mauvaises réponses, quelles ont été leurs erreurs ? Dans 82% des cas, ils ont fait une erreur de suffixe dérivationnel, choisissant le compétiteur le plus fréquent de la même famille. Cette nouvelle observation vient nous éclairer quant à la stratégie du « ne sait pas » et de sa réelle signification. Dans le cas des participants du groupe LPA, le « ne sait pas » signifie en réalité qu'ils ont fait une recherche parmi les mots qu'ils connaissaient et qu'ils ont tout simplement écrit celui qui leur est venu en tête en premier, sans se poser d'autres questions. Comme l'ont suggéré Feldman et Stemmer pour expliquer ce type de réponse, ils n'ont probablement pas procédé à l'aide d'un raisonnement verbal et voilà ce qui explique pourquoi ils n'ont pas pu verbaliser sur la manière dont ils avaient complété les mots. Le « ne sait pas », dans cette situation, est donc en fait la stratégie de la recherche dans le lexique mental mais déguisée.

Cette dernière observation nous ramène à notre hypothèse que les participants du groupe LPA ont tout simplement des connaissances lexicales plus pauvres que les

participants du groupe LBA et que c'est pour cette raison, et non pas parce qu'ils ne sont pas capables d'analyser, qu'ils ont moins bien performé lors du traitement des mots dérivés dans nos C-tests.

Voyons maintenant si nous pouvons obtenir une évidence similaire avec l'analyse des correspondances entre stratégies et réponses.

3.2.3.3 Correspondance entre stratégies et réponses au 3^e C-test

Nous avons jugé qu'il serait intéressant de vérifier à quel point il y a une correspondance entre le type de stratégie choisi et le mot écrit dans le C-test. Pour ce faire, nous avons écouté à nouveau les entrevues et classé chaque mot interviewé pour chaque participant selon la grille suivante:

Tableau 3.15
Codage de la correspondance entre stratégie et réponse

Type de correspondance	Légende
Bonne stratégie nommée par le sujet et bonne réponse	(S-R)
Bonne stratégie nommée par le sujet mais mauvaise réponse	(S-X)
Mauvaise stratégie nommée par le sujet mais bonne réponse	(X-R)
Mauvaise stratégie nommée par le sujet et mauvaise réponse	(X-X)

Pour ce qui est de la bonne ou de la mauvaise réponse, il nous suffisait de vérifier si le mot choisi par le participant était le bon et s'il avait été écrit correctement. Par contre, pour ce qui est de différencier la bonne de la mauvaise stratégie, nous avons compilé, pour chaque mot interviewé, les stratégies qui permettaient effectivement d'obtenir des indices pertinents pour trouver le mot en question. Toutes les autres stratégies étaient alors classées comme mauvaises. Voici un exemple :

Exemple 1 (Participant 306)

- Extrait du C-test + réponse écrite par le participant en italique
Efficacité : Travailler *efficacitive*, c'est avoir (...)
- Extrait de l'entrevue de la tâche de rappel

<u>Interviewer :</u>	Très bien. Maintenant, regardons le deuxième paragraphe. Comment avez-vous trouvé ce mot ? (L'interviewer pointe du doigt le mot <i>efficacitive</i> .)
<u>Participant :</u>	C'est un adverbe mais je ne trouve pas quels mots mettre à la fin. Mais dans ma tête, je n'ai pas les mots souvent.

Dans cet extrait, on remarque que le participant a identifié et utilisé la nature du mot comme indice. Nous avons jugé cette stratégie bonne car elle aurait pu permettre au participant de trouver le suffixe approprié pour compléter le mot. Malheureusement, ce participant n'a pas réussi à trouver le bon mot. Nous avons donc classé cet exemple dans la catégorie : bonne stratégie mais mauvaise réponse (S-X). Voici un deuxième exemple :

Exemple 2 (Participant 102)

- Extrait du C-test + réponse écrite par le participant en italique :
Plusieurs facteurs peuvent être une source de stress pour des étudiants adultes : (...) les pressions *parentales* (...)
- Extrait de l'entrevue de la tâche de rappel :

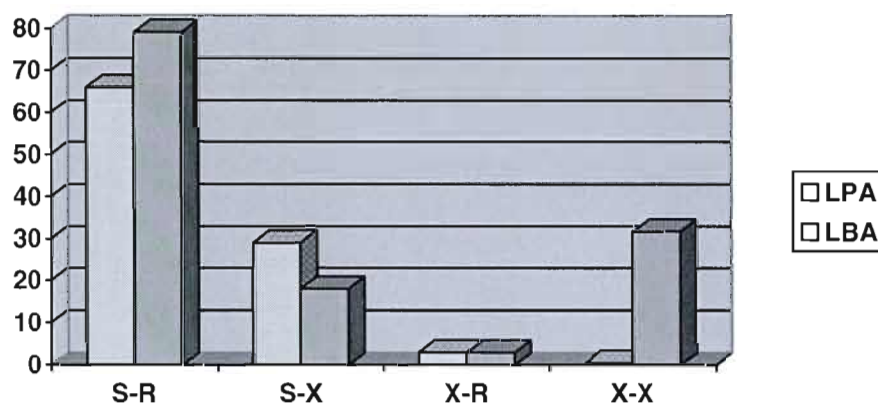
<u>Interviewer :</u>	Comment avez-vous trouvé ce mot ? (L'interviewer pointe le mot <i>parentales</i> .)
<u>Participant :</u>	Je suis aussi étudiant (le participant pointe le mot <i>étudiant</i>) et je pense qu'avec le sens ..., ça allait avec parent alors j'ai pensé <i>pressions parentales</i> .

Le participant de cet extrait mentionne deux stratégies que nous avons jugées bonnes pour compléter le mot tronqué. Il a mentionné tout d'abord le sens de la phrase étant donné qu'il a fait référence au mot *étudiant* écrit au début de la phrase; et puis il a mentionné la morphologie dérivationnelle en parlant de la base du mot qui est, disons-le, une base très transparente. Cette base lui a permis de définir tout de suite la famille de mots dans laquelle il devait sélectionner le mot

tronqué. Finalement, pour ce qui est du mot en question, le participant l'a écrit correctement. Pour ces raisons, nous avons classé ce mot dans la catégorie : bonne stratégie et bonne réponse (S-R).

Passons maintenant à la présentation de la compilation de nos résultats :

Figure 3.5
Correspondance entre les stratégies identifiées lors de la tâche de
rappel stimulé et les réponses écrites dans le 3^e C-test



Comme nous pouvons le constater, et c'est une bonne nouvelle, en général et ce, même si les participants du LBA ont obtenu plus souvent de bons résultats que leurs camarades de classe, les participants des deux groupes connaissent de bonnes stratégies et celles-ci leur ont permis de trouver les bonnes réponses. La différence significative entre ces deux groupes au niveau de la réussite globale et de la réussite des mots dérivés ne consiste donc pas en des erreurs au niveau des stratégies.

Dans une proportion tout de même faible, on constate que les participants du groupe LPA sont ceux qui ont le plus souvent écrit la mauvaise réponse et ce, même s'ils avaient de bonnes stratégies. Nous avons peut-être ici dans cette observation la réponse à notre hypothèse. Si les participants du groupe LPA ont

souvent de bonnes stratégies qui les mènent aux bonnes réponses, ils leur arrivent également plus souvent que les autres d'avoir de bonnes stratégies mais d'écrire de mauvaises réponses. Pourquoi ces erreurs ? La réponse à cette question nous a été donnée par les analyses de nos 3 C-tests. Revenons en arrière.

Lors de l'analyse des résultats des deux premières épreuves, nous avons pu observer que la différence significative au niveau de la réussite des mots dérivés entre les deux groupes était principalement causée par des erreurs de suffixes, surtout lorsque les mots étaient peu fréquents ou que leur suffixe était peu fréquent. Avec les résultats des analyses du 3^e C-test, nous nous sommes aperçus que ce n'était pas la fréquence des suffixes qui compliquaient le traitement des mots dérivés pour les participants du groupe LPA, mais bien plutôt la fréquence même des mots. Ils ne pouvaient peut-être pas trouver le nom *difficulté*, mais connaissaient bien l'adjectif *difficile* soit un mot de la même famille mais beaucoup plus fréquent.

Les participants du groupe LPA feraient donc plus d'erreurs dans le traitement des mots dérivés en français non pas parce qu'ils ne connaissent pas de bonnes stratégies pour décoder et produire des mots, mais tout simplement parce qu'ils ne connaissent pas assez de mots et ne peuvent donc pas choisir celui qui correspond le plus aux indices fournis par le texte.

Pour faire un bref résumé, les deux groupes connaissent en général les mêmes stratégies et les utilisent efficacement pour décoder et compléter des mots dérivés. Nous avons toutefois noté de petites différences. Les participants du groupe LBA se sont référés plus souvent au sens de la phrase et à la fonction du mot pour compléter les mots tronqués alors que les participants du groupe LPA ont plus souvent eu recours à la recherche dans leur lexique mental. Cette stratégie peut toutefois être problématique si le sujet n'a pas un vocabulaire très riche. Résultat : il écrit la mauvaise réponse non pas par choix, mais plutôt parce que le vocabulaire fait défaut. Cette hypothèse est appuyée par les stratégies

discriminant les deux groupes ainsi que par les correspondances que nous avons trouvées entre les stratégies et les réponses.

3.3 Synthèse

Nous ferons tout d'abord un retour sur les diverses observations que nous avons faites suite aux analyses des C-tests (3.3.1) et des stratégies mentionnées lors de la tâche de rappel stimulé (3.3.2). Nous terminerons en vérifiant dans quelle mesure ces découvertes peuvent confirmer ou, à tout le moins, éclairer nos hypothèses de recherche (3.3.3).

3.3.1 Résumé – Résultats des C-tests

En général, les participants du groupe LBA ont obtenu de meilleurs résultats globaux ainsi que de meilleurs résultats pour le traitement des mots dérivés que leurs camarades de classe appartenant au groupe LPA. La différence entre leurs résultats était statistiquement significative (sauf pour le troisième C-test).

Quant aux participants du groupe LPA, s'ils se sont améliorés au niveau des résultats globaux (familiarisation avec l'exercice ; effet de nouveauté dissipé), ils ont néanmoins continué d'éprouver des problèmes avec le traitement des mots dérivés. Que ce soit lors du premier, du deuxième ou du troisième C-test, la différence au niveau du décodage et de la production des mots dérivés entre les deux groupes est demeurée statistiquement significative. Ces résultats, mentionnons-le, correspondent aux résultats obtenus en anglais langue seconde à savoir que plus la L1 est distante du point de vue morphologique de la L2, et plus l'analyse des mots de la L2 est difficile.

Nous avons voulu ensuite approfondir notre réflexion en cherchant à connaître la cause d'une telle difficulté. Pourquoi les participants du groupe LPA ne peuvent-ils pas traiter les mots dérivés avec autant de succès que leurs camarades de classe? Pour tenter de répondre à cette question, nous avons effectué plusieurs analyses au niveau même des mots dérivés afin de connaître, entre autres, la nature des mots qui différenciaient les deux groupes. Dans le cas du premier et du deuxième C-tests, nous avons découvert que les mots discriminant les deux groupes étaient soit des mots peu fréquents dans le vocabulaire, soit ils avaient été construits à l'aide de suffixes peu fréquents. Pour ce qui est du 3^e C-test, les mots discriminants semblaient tout simplement ne pas être connus par les participants du groupe LPA.

3.3.2 Résumé - Résultats de la tâche de rappel stimulé

Nous nous sommes ensuite intéressés aux stratégies qui avaient été utilisées par les participants des deux groupes pour compléter les mots dérivés. Afin de pouvoir dresser un portrait assez global des stratégies utilisées, nous avons d'abord regroupé celles-ci en quatre catégories. En procédant ainsi, nous nous sommes rendus compte que les participants des deux groupes utilisaient sensiblement les mêmes catégories de stratégies dans les mêmes proportions. Entre autres, peu importe les caractéristiques morphologiques du système d'écriture de leur L1, ils se sont tous fiés en premier au contexte immédiat (les premières lettres du mot, le nombre de tirets, le mot avant et celui après, etc.) pour compléter les mots tronqués. (Ce qui rappelle les résultats obtenus par Koda (1999) dans une étude sur la reconnaissance de mots en anglais.)

Si les participants ont tous utilisé les mêmes catégories de stratégies, nous avons toutefois constaté de petites différences quand nous avons examiné chaque stratégie individuellement. Alors que les participants du groupe LBA se sont fiés davantage au sens de la phrase et à la fonction du mot pour traiter les mots

dérivés, les participants du groupe LPA, eux, se sont référés plus souvent à leur lexique mental. Cette stratégie devient problématique si le sujet a peu de connaissances lexicales. Étant donné la différence significative au niveau de la réussite des mots dérivés entre les deux groupes, nous pouvons nous imaginer que les participants n'avaient tout simplement pas le bagage lexical pour compléter le C-test. Cette supposition est appuyée par l'analyse de la correspondance entre les stratégies identifiées par les participants et leurs réponses écrites lors du 3^e C-test. Alors que dans la majorité des cas, les participants des deux groupes ont utilisé de bonnes stratégies qui les ont aidés à trouver les bonnes réponses, les participants du groupe LPA ont plus souvent écrit de mauvaises réponses tout en s'étant fiés à de bonnes stratégies. Les participants du groupe LPA n'éprouvent donc pas de problème au niveau des stratégies qu'ils emploient, mais plutôt au niveau des mots qu'ils connaissent. Ils n'en connaissent tout simplement pas assez. Ceci nous amène à émettre la supposition que les participants du groupe LPA ont peut-être un vocabulaire plus pauvre que les participants du groupe LBA.

3.3.3 Synthèse – Hypothèses de recherche

Dans notre chapitre sur la méthodologie, nous avons présenté les deux hypothèses de recherche que voici :

Étant donné que le français est une langue écrite riche en affixes :

- 1- La performance au niveau du décodage et de la production de mots suffixés en français langue seconde sera moins bonne chez les apprenants dont la langue maternelle écrite a peu d'affixes que chez les apprenants dont la langue maternelle écrite comprend davantage affixes.
- 1- Les stratégies employées au niveau du décodage et de la production de mots suffixés en français langue seconde par des apprenants dont la langue

maternelle écrite contient peu d'affixes seront différentes de celles qui sont employées par des apprenants dont la langue maternelle écrite comprend davantage d'affixes.

Suite à nos résultats, que pouvons-nous dire de ces deux hypothèses de recherche?

Hypothèse 1 :

Les résultats des diverses analyses que nous avons faites à l'aide des données recueillies lors de nos trois C-tests nous suggèrent que notre première hypothèse est exacte. Effectivement, la performance au niveau du traitement des mots suffixés en français langue seconde est moins bonne chez des apprenants dont la langue maternelle écrite a peu d'affixes que chez des apprenants dont la langue maternelle écrite comprend davantage d'affixes. Non seulement nous avons constaté une différence importante entre leurs moyennes, mais cette différence est statistiquement significative pour les trois C-tests. Cette différence, comme nous l'avons largement expliquée lors de la synthèse des résultats des C-tests et des entrevues, ne serait pas causée par un manque de connaissances des procédés de dérivation en français, mais plutôt par un manque de connaissances lexicales.

Hypothèse 2 :

Suite aux observations que nous avons recueillies grâce à la compilation des diverses stratégies identifiées par les participants, nous ne pouvons pas entièrement confirmer notre hypothèse. S'il existe de petites différences au niveau des stratégies nommées par les deux groupes lorsque celles-ci sont examinées individuellement, ces différences disparaissent lorsque nous regroupons les stratégies. En fait, les participants des deux groupes utilisent les mêmes stratégies et, dans une très grande proportion, les utilisent efficacement. Ce que nous avons pu observer, toutefois, c'est que les participants dont la langue première écrite contient peu d'affixes s'étaient fiés davantage à leur lexique mental pour compléter les mots tronqués et cela avait eu un effet fâcheux. Les

résultats nous permettent d'avancer l'hypothèse que ce sont encore une fois les connaissances lexicales de ces participants qui sont en cause.

CONCLUSION

De nos jours, les défis sont nombreux pour les enseignants du français langue seconde qui doivent constamment ajuster leur enseignement à une clientèle de plus en plus diversifiée culturellement et linguistiquement. Parmi ces défis, on retrouve celui posé par la question du transfert des connaissances de la L1 vers la L2, surtout lorsqu'il est question d'étudiants adultes. Dans le cadre de ce mémoire, nous nous sommes intéressés à l'influence des caractéristiques morphologiques du système d'écriture de la langue première sur la maîtrise de la dérivation du français. Nous avons voulu vérifier si ces caractéristiques morphologiques avaient un impact sur a) le traitement (décodage et production) de mots suffixés et sur b) les stratégies employées pour compléter des mots suffixés tronqués.

Afin d'être en mesure de répondre à ces questions de recherche, nous avons administré deux types d'épreuves auprès d'une soixantaine d'étudiants adultes de niveau universitaire provenant de divers horizons linguistiques. Nous leur avons d'abord demandé de compléter deux C-tests à sept semaines d'intervalle. Par la suite, tous ceux qui désiraient poursuivre l'expérience ont été conviés à une tâche de rappel stimulé.

Les résultats que nous avons obtenus grâce à nos épreuves nous ont permis de faire plusieurs découvertes intéressantes dont, entre autres, celle de l'influence non négligeable des caractéristiques morphologiques du système d'écriture de la langue première sur le traitement des mots dérivés. Nous avons constaté que les participants qui avaient une langue première écrite pauvre en affixes (LPA) avaient de piètres résultats au niveau du décodage et de la production de mots dérivés comparativement à leurs camarades de classe ayant une L1 écrite riche en affixes et ce, peu importe le nombre de fois qu'ils ont eu à compléter des C-tests.

Cette contre-performance, statistiquement significative, ne remet pas en cause les capacités d'analyse ni les stratégies déployées par ces étudiants pour traiter les mots dérivés. Bien au contraire, ces participants connaissent et utilisent les mêmes stratégies que leurs camarades de classe et se réfèrent à l'analyse grammaticale et lexicale aussi souvent qu'eux. La cause de cette incapacité à traiter les mots dérivés résiderait fort probablement dans un manque de connaissances lexicales.

Si les participants qui ont une L1 écrite pauvre en affixes font autant d'erreurs de suffixes que les autres lorsque les mots à compléter sont peu fréquents dans leur vocabulaire (ou ont un suffixe peu fréquent), ils continuent néanmoins à faire autant d'erreurs lorsque les mots à compléter sont fréquents (ou ont un suffixe fréquent). Dans la majorité des cas, ils choisissent le premier mot qui leur vient en tête et qui est souvent un mot compétiteur de la même famille et ce, même s'ils savent à l'avance que ce mot n'est pas le bon. Cette stratégie de la recherche du mot dans le lexique mental peut engendrer de fâcheux résultats lorsqu'il y a peu de mots disponibles dans le lexique mental.

Comme nous l'avons observé lors de la compilation des stratégies identifiées par les participants pour compléter les mots tronqués du C-test, la stratégie de la recherche dans le lexique mental est une stratégie qui a été beaucoup plus souvent mentionnée par les participants du groupe LPA que les autres. Qui est plus, dans la plupart des cas, cette stratégie a été associée à de mauvaises réponses lors de l'épreuve. Ce constat suggère que les participants du groupe LPA n'ont tout simplement pas accès à un vocabulaire assez riche.

Selon nous, pour aider ces étudiants à améliorer leurs résultats, il faudrait avant tout les amener à diversifier et à accroître leurs connaissances lexicales. Une bonne manière de procéder serait d'enseigner systématiquement les suffixes, leur rôle et leur sens en classe car, comme l'expliquent Schmitt et Meara (1997), les suffixes et le vocabulaire sont étroitement liés :

« Greater knowledge of suffixes would coincide with a larger vocabulary because it would facilitate access to more members of a word's family. This larger vocabulary size would make more words available for associations. » (p.20)

Si l'enseignement d'une telle stratégie présente des écueils (Clark et Nation, 1980; Carlisle, 1988), elle offre néanmoins l'avantage d'être une stratégie facile à enseigner et qui requiert peu de temps de classe. L'étudiant peut ainsi augmenter rapidement et ce, avec peu d'efforts cognitifs, ses connaissances lexicales. Dans le cas de nos participants qui ont souvent choisi un compétiteur de la même famille pour compléter des mots pourtant fréquents, l'enseignement d'une telle stratégie leur permettrait d'élargir leurs connaissances des champs lexicaux et des familles de mots.

Nulle recherche n'est parfaite et la nôtre n'échappe pas à cette règle. Parmi ses limites, mentionnons d'abord l'absence d'un test vérifiant les connaissances lexicales des participants. Puisque nous pensions que le problème résidait au niveau des connaissances de la dérivation en français et que les difficultés surviendraient surtout avec les mots ayant un suffixe peu fréquent, nous n'avons pas planifié l'administration d'une troisième épreuve, en l'occurrence ici un test de vérification des connaissances lexicales. Ce n'est qu'une fois les résultats analysés et l'hypothèse de la pauvreté lexicale soulevée que nous avons considéré l'ajout de ce type de test. Cependant, étant déjà parvenu à la fin de la session, il nous était à ce moment-là impossible d'administrer un nouveau test aux mêmes étudiants.

Une autre limite de cette recherche a été le nombre assez faible de participants volontaires pour la tâche de rappel stimulé. Il aurait été intéressant d'évaluer un échantillon plus grand de participants pour vérifier les tendances que nous avons observées au niveau des stratégies employées lors des C-tests. Puisque notre tâche de rappel stimulé s'est déroulée sur une base volontaire, trop peu de sujets

ont été évalués pour justifier des analyses statistiques qui nous auraient permis de confirmer nos observations.

Finalement, mentionnons que selon notre modèle théorique, le BIA+, les représentations orthographiques ne sont pas les seules à activer le lexique mental, bien qu'elles aient le rôle principal (Pexman et Lupker, 2001). Les représentations phonologiques et sémantiques interagissent avec les représentations orthographiques. Puisque notre recherche a porté sur la reconnaissance de mots écrits, c'est-à-dire de représentations orthographiques, il se pourrait que la clé de l'accès au lexique mental des participants du groupe LPA se trouve plutôt du côté de la phonologie et de la sémantique. Deux autres recherches complémentaires, une portant sur l'aspect phonologique et l'autre sur l'aspect sémantique, auraient alors été nécessaires.

Si notre recherche a ses limites, elle offre également de nouvelles pistes de réflexion pour de futures recherches. Il serait intéressant, entre autres, de refaire la même expérience mais auprès d'une clientèle ayant une langue première dite « isolante » autre que le mandarin. Dans notre étude, tous les participants du groupe LPA parlaient le mandarin. Les résultats auraient-ils été les mêmes avec un autre groupe d'apprenants ?

Comme nous l'avons mentionné plus tôt, dans une future recherche, il serait également pertinent d'ajouter un test de connaissances lexicales en plus des C-tests et de la tâche de rappel stimulé afin de pouvoir statistiquement vérifier s'il y a bel et bien un lien étroit entre les résultats obtenus lors du décodage et de la production de mots dérivés et les résultats obtenus lors du test d'évaluation des connaissances lexicales.

Finalement, il serait instructif de vérifier l'impact de l'enseignement systématique des suffixes sur les résultats obtenus au niveau des mots dérivés d'un C-test. Une telle recherche pourrait se dérouler en trois étapes avec d'abord un prétest, soit ici

en l'occurrence un premier C-test; une période durant laquelle il y aurait enseignement systématique des principaux suffixes et introduction aux champs lexicaux; et finalement, dans un troisième temps, l'administration d'un 2^e C-test, en post-test, permettant ainsi de vérifier si les participants ont amélioré leurs résultats au niveau du traitement des mots dérivés. Ce type de recherche pourrait, entre autres, jeter un nouvel éclairage sur les connaissances lexicales des apprenants ayant une L1 écrite pauvre en affixes.

Il reste encore bien du chemin à faire pour comprendre comment des apprenants d'horizons linguistiques différents font pour maîtriser la dérivation en français. Si nous avons pu, grâce à notre recherche, apporter des pistes de solution, il n'en demeure pas moins que la dérivation est un aspect encore peu étudié en français langue seconde et qui mériterait pourtant, à bien des égards, ne serait-ce que pour son lien étroit avec le vocabulaire, que l'on s'y attarde davantage.

APPENDICE A

C-TESTS

Date: _____ Nom: _____ Groupe: _____
 Sexe: M ☐ F ☐ Langue maternelle : _____

Consigne : Vous avez 15 minutes pour compléter le texte suivant.

Le retour aux études - Service à la vie étudiante

Le retour aux études se caractérise par de nombreuses situations d'adaptation.

En pl_ _ de dev_ _ faire fa_ _ aux exige_ _ _ nouve_ _ _ , ceux qu_ _
 reviennent a_ _ études apr_ _ une certaine abse_ _ _ sont confro_ _ _ à
 tou_ _ sortes de diffic_ _ _ en rais_ _ de la méconna_ _ _ du
 mil_ _ et du man_ _ de « prat_ _ » face a_ _ tâches scol_ _ _ à
 effec_ _ _ . Plusieurs so_ _ tentés d'aband_ _ _ dès l_ _ premières
 sema_ _ _ à ca_ _ de la pres_ _ _ qu'ils resse_ _ _ et l'impre_ _ _ de
 n_ pas av_ _ les capa_ _ _ intellec_ _ _ pour répo_ _ _ aux
 exige_ _ _ .

Afin d_ mieux surm_ _ _ les diffic_ _ _ qui so_ _ reliées au ret_ _ _ aux
 étu_ _ , la sect_ _ _ « Aide et sout_ _ _ à l'appren_ _ _ » offre aux
 étud_ _ _ une confé_ _ _ intitulée « **Appriv** _ _ _ **le retour aux**
études ». La présen_ _ _ comprend:

- une descri_ _ _ du déroul_ _ _ d'une prem_ _ _ session _
 l'université av_ _ les pha_ _ _ qu'elle comp_ _ _ et les diffic_ _ _
 particulières q_ _ la caracté_ _ _ ;
- les règ_ _ _ « informelles » à resp_ _ _ pour facil_ _ _ son
 adapt_ _ _ et son intég_ _ _ ;
- les imp_ _ _ sur l_ vie person_ _ _ , famil_ _ _ ou soci_ _ _
- des explic_ _ _ des phéno_ _ _ d'adapt_ _ _ qui sont
 inévit_ _ _ ;
- plusieurs solu_ _ _ ou sugge_ _ _ pour ai_ _ _ à persé_ _ _ .

L'atel_ _ _ est d'une dur_ _ de 90 min_ _ _ et ne néces_ _ _ **aucune**
inscri _ _ **préa** _ _ . Il suf_ _ _ de vo_ _ présenter à l'u_ _ des renc_ _
 _ _ _ qui vo_ _ sont offe_ _ _ , au loc_ _ et _ l'heure correspo_ _ _ à
 vos disponi_ _ _ .

Un moy_ _ efficace d_ s'outi_ _ _ pour mi_ _ _ affronter l'adapt_ _ _ aux
 études univers_ _ _ !

Date: _____ Nom: _____ Groupe: _____
 Sexe: M ☐ F ☐ Langue maternelle : _____

Consigne : Vous avez 15 minutes pour compléter le texte suivant.

Profil général de l'immigration au Québec

(Montréal : Laboratoire de cosmopolitisme entre deux mondes Annick Germain,
 Juin 1997)

L'immigration au Canada est un phénomène en pleine croissance. En ef_ _ _ , le Can_ _ _ est u_ _ des rar_ _ pays industr_ _ _ _ _ à accue_ _ _ _ _ une immigr_ _ _ _ _ interna_ _ _ _ _ importante. A_ _ Québec seul_ _ _ _ _ , 1 100 000 immigr_ _ _ _ _ ont é_ _ _ accueillis d_ _ 1946 _ 1986.

Le Qué_ _ _ , tout com_ _ le Canada, sélect_ _ _ _ _ ses immigr_ _ _ _ _ et diver_ _ _ _ _ les catég_ _ _ _ _ d'admis_ _ _ _ _ . La propo_ _ _ _ _ d'immigrants ad_ _ _ au Québec comme indépe_ _ _ _ _ (sélectionnés sel_ _ leur niv_ _ _ de scol_ _ _ _ , leur expe_ _ _ _ _ professi_ _ _ _ _ , leur â_ _ , leur connai_ _ _ _ _ lingui_ _ _ _ _ et les beso_ _ _ du mar_ _ _ de l'emp_ _ _) de 1992 à 1996 est de 44,9 %; cel_ _ _ des personnes admi_ _ _ sous le sta_ _ _ de parent aidé e_ _ _ de 2,4 %; le statut d_ _ réunion fami_ _ _ _ _ regroupe 34,2 % d_ _ immigrants adm_ _ _ au Qué_ _ _ et les réfu_ _ _ _ _ représentent 20,9 % d_ _ immigrants.

En 1991, 88 % de la popul_ _ _ _ _ immi_ _ _ _ _ au Québec rési_ _ _ _ _ dans l_ _ région métropo_ _ _ _ _ de Mont_ _ _ _ _ . Cette concent_ _ _ _ _ de l'immig_ _ _ _ _ dans l_ _ seule rég_ _ _ montré_ _ _ _ _ a d_ _ _ conséquences tr_ _ importantes _ de multi_ _ _ _ _ niveaux. Par consé_ _ _ _ _ , la ques_ _ _ _ _ de l'immi_ _ _ _ _ est u_ _ des élém_ _ _ _ _ qui disti_ _ _ _ _ la région métropo_ _ _ _ _ du res_ _ _ du Qu_ _ _ _ _ .

Par con_ _ _ , Montréal com_ _ _ moins _ ' immigrants q_ _ _ les aut_ _ _ villes canad_ _ _ _ _ . Depuis 1991, le gouver_ _ _ _ _ du Québec ten_ _ _ d'encou_ _ _ _ _ l'établi_ _ _ _ _ dans les aut_ _ _ régions par dive_ _ _ _ _ mesures incita_ _ _ _ _ . Mais il sem_ _ _ bien diff_ _ _ _ _ d'inve_ _ _ _ _ des tend_ _ _ _ _ à la concent_ _ _ _ _ dans les parc_ _ _ _ _ migra_ _ _ _ _ .

Date: _____ Nom: _____ Groupe: _____
 Sexe: M ☐ F ☐ Langue maternelle : _____

Consigne : Vous avez 15 minutes pour compléter le texte suivant.

La gestion du temps - Service à la vie étudiante

Pour être capable de planifier un bon horaire et de le respecter, il faut tenir compte de trois notions de base : l'équilibre, l'efficacité et le perfectionnisme.

A) Équilibre : Plusieurs fact_ _ _ _ peuvent êt_ _ une sou_ _ _ de str_ _ _ pour des étudi _ _ _ _ adultes: les échéan_ _ _ _ très rapp_ _ _ _ , les exige _ _ _ _ du profe_ _ _ _ , le continge_ _ _ _ des progr_ _ _ _ , les soucis profess_ _ _ _ , les problèmes finan_ _ _ _ , les ruptures amour_ _ _ _ , les press_ _ _ _ paren_ _ _ _ , etc. On les appe_ _ _ des stres_ _ _ _ externes. Le stress cau_ _ par cert_ _ _ _ facteurs pe_ _ être réd_ _ _ si on équil_ _ _ son hora_ _ _ et p_ _ seulement s_ _ horaire scol_ _ _ ! En fa_ _ , quand i_ est ques_ _ _ _ d'équil_ _ _ , il s'ag_ _ de la répart_ _ _ _ équit_ _ _ _ de son tem_ _ entre les étu_ _ , le trav_ _ , la fami_ _ , etc.

B) Efficacité : Travailler effica_ _ _ _ , c'est avo_ _ la capa_ _ _ _ de prod_ _ _ le maxi_ _ _ de résul_ _ _ avec un mini_ _ _ d'effo_ _ _ . Un horaire surch_ _ _ _ peut vo_ _ rendre ineff_ _ _ _ car ri_ _ ne sert d'étud_ _ _ quand vo_ _ n'êtes pl_ _ concentré. La concent_ _ _ _ varie sel_ _ la diffi_ _ _ _ de la tâc_ _ , la réfle_ _ _ _ qu'el_ _ vous dema_ _ _ ou tout simpl_ _ _ _ la dur_ _ de la concent_ _ _ _ exigée. De faç_ _ générale, lo_ _ de la planifi_ _ _ _ de votre hora_ _ , dix minu_ _ _ de pau_ _ sont nécess_ _ _ _ pour cha_ _ _ heure de travail intell_ _ _ _ .

C) Perfectionnisme : La quêt_ _ de l'excel_ _ _ _ et de la perfe_ _ _ _ mène souv_ _ _ les indiv_ _ _ _ à se fix_ _ _ des stand_ _ _ _ de réus_ _ _ trop éle_ _ _ . Ceci pe_ _ s'avérer démot_ _ _ _ et avo_ _ comme résu_ _ _ la procrast_ _ _ _ . Il est do_ _ important de se fix_ _ _ des atte_ _ _ _ réali_ _ _ _ .

Liste des mots suffixés sélectionnés pour le C-test 1

Mot dérivé	% de réussite (LPA)	% de réussite (LBA)	Catégorie grammaticale	Fréquence* TFF \ TFL	Productivité**
<i>Exigences (1)***</i>	25	78	Nom	4.76 \ 17.36	430
<i>Absence</i>	25	78	Nom	23.31 \ 76.28	430
<i>Difficultés (1)</i>	53	93	Nom	17.11 \ 49.80	900
Méconnaissance	47	52	Nom	0 \ 0.81	430
Scolaires	67	78	Adjectif	8.13 \ 14.12	600
<i>Pression</i>	33	63	Nom	40.30 \ 21.76	1900
Capacités	53	74	Nom	17.89 \ 14.39	900
<i>Exigences (2)</i>	28	74	Nom	4.76 \ 17.36	430
<i>Difficultés (2)</i>	53	96	Nom	17.11 \ 49.80	900
Apprentissage	28	41	Nom	2.65 \ 10.47	1000
Étudiants	81	85	Nom	37.41 \ 38.45	1400
Présentation	67	81	Nom	9.76 \ 8.38	1900
Description	81	89	Nom	6.57 \ 11.49	1900
<i>Déroulement</i>	44	78	Nom	1.45 \ 5.95	1000
<i>Difficultés (3)</i>	56	93	Nom	17.11 \ 49.80	900
Adaptation (1)	75	93	Nom	1.81 \ 4.66	1900
<i>Intégration</i>	53	93	Nom	0.96 \ 1.08	1900
Familiale	56	56	Adjectif	12.77 \ 32.84	800
Sociale	63	63	Adjectif	24.76 \ 49.86	800
Adaptation (2)	89	89	Nom	1.81 \ 4.66	1900
<i>Inévitables</i>	67	67	Adjectif	10.72 \ 16.49	800
Suggestions	63	63	Nom	11.14 \ 7.77	1900
<i>Durée</i>	93	93	Nom	8.01 \ 19.39	400
Inscription	93	93	Nom	7.89 \ 19.32	1900
Rencontres	52	52	Nom	28.86 \ 81.76	n/a

<i>Disponibilités</i>	56	56	Nom	0.66 \ 3.58	900
-----------------------	----	----	-----	-------------	-----

* Ces taux de fréquence ont été tirés du site « www.lexique.org », subventionné par le CNRS. Les taux de fréquence présentés ici (TFF : taux de fréquence films; TFL : taux de fréquence livres) ont été calculés selon la formule suivante : la somme des fréquences du lemme et de ses formes fléchies sur 1 million d'occurrences. Pour de plus amples renseignements sur le calcul de ces taux, se référer au site.

** Productivité compilée par Catherine-Marie Longtin (2004).

*** D'après des analyses ANOVA, tous les mots en caractères gras sont des mots discriminant les deux groupes (la différence au niveau du pourcentage de réussite de ces mots étant statistiquement significatives).

Liste des mots suffixés sélectionnés pour le C-test 2

Mot dérivé	% de réussite (LPA)	% de réussite (LBA)	Catégorie grammaticale	Fréquence* TFF \ TFL	Productivité**
<i>Industrialisés***</i>	6	33	Adjectif	0.06 \ 0.14	450
Immigration (1)	78	89	Nom	4.28 \ 0.81	1900
Internationale	69	70	Adjectif	14.04 \ 24.86	800
Immigrants (1)	94	96	Nom	1.33 \ 0.88	1400
Immigrants (2)	92	96	Nom	1.33 \ 0.88	1400
Diversifier	11	22	Verbe	1.02 \ 0.88	95
<i>Catégories</i>	44	93	Nom	7.53 \ 13.24	1300
Admission	67	70	Nom	4.46 \ 1.82	1900
<i>Proportion</i>	11	56	Nom	3.07 \ 12.70	1900
Indépendants	33	33	Nom	1.27 \ 0.61	1400
<i>Scolarité</i>	28	81	Nom	1.27 \ 0.81	900
Expertise	3	11	Nom	2.65 \ 0.68	30
<i>Connaissance</i>	94	78	Nom	43.73 \ 67.50	430
<i>Linguistique</i>	64	96	Adjectif	0.12 \ 1.55	1600
Familiale	39	41	Adjectif	12.77 \ 32.84	800
Réfugiés	8	22	Nom	7.05 \ 10.27	400
Population	89	96	Nom	12.65 \ 33.65	1900
Immigrée	36	52	Adjectif	0.24 \ 1.28	400
Métropolitaine (1)	44	67	Adjectif	0.78 \ 3.92	90
<i>Concentration (1)</i>	53	85	Nom	7.53 \ 10.41	1900
Immigration (2)	78	59	Nom	4.28 \ 0.81	1900
Montréalaise	64	70	Adjectif	n/a	35
<i>Multiples</i>	17	56	Adjectif	6.08 \ 18.18	n/a
Immigration (3)	83	93	Nom	4.28 \ 0.81	1900
Métropolitaine (2)	42	63	Adjectif	0.78 \ 3.92	90
Canadiennes	69	74	Adjectif	4.58 \ 4.93	520

Établissement	47	44	Nom	6.08 \ 22.23	1000
<i>Diverses</i>	25	56	Adjectif	6.69 \ 52.84	n/a
<i>Incitatives</i>	0	19	Adjectif	n/a	375
Tendances	22	33	Nom	14.52 \ 20.88	430
Concentration (2)	50	70	Nom	7.53 \ 10.41	1900
<i>Migratoires</i>	0	41	Adjectif	n/a	160

* Ces taux de fréquence ont été tirés du site « www.lexique.org », subventionné par le CNRS. Les taux de fréquence présentés ici (TFF : taux de fréquence films; TFL : taux de fréquence livres) ont été calculés selon la formule suivante : la somme des fréquences du lemme et de ses formes fléchies sur 1 million d'occurrences. Pour de plus amples renseignements sur le calcul de ces taux, se référer au site.

** Productivité compilée par Catherine-Marie Longtin (2004).

*** D'après des analyses ANOVA, tous les mots en caractères gras sont des mots discriminant les deux groupes (la différence au niveau du pourcentage de réussite de ces mots étant statistiquement significatives).

Liste des mots suffixés sélectionnés pour le C-test 3

Mot dérivé	% de réussite (LPA)	% de réussite (LBA)	Catégorie grammaticale	Fréquence* TFF \ TFL	Productivité**
Exigences	33	67	Nom	4.76 \ 17.36	430
Contingement	0	0	Nom	0.06 \ 0	1000
Échéanciers	0	0	Nom	0 \ 0.14	800
Financiers	25	58	Adjectif	10.84 \ 9.39	800
Pressions	67	92	Nom	40.30 \ 21.76	1900
<i>Parentales****</i>	8	58	Adjectif	1.51 \ 0.81	800
Stresseurs	0	8	Nom	n/a	1500
Scolaire	92	83	Adjectif	8.13 \ 14.12	600
<i>Répartition</i>	0	50	Nom	0.90 \ 3.11	1900
Équitable	8	42	Adjectif	4.04 \ 2.97	800
<i>Efficacement</i>	33	83	Adverbe	1.08 \ 2.23	1350
Capacité	92	100	Nom	17.89 \ 14.39	900
Résultats	58	67	Nom	52.71 \ 38.38	175
<i>Concentration (I)</i>	67	100	Nom	7.53 \ 10.41	1900
Difficulté	75	75	Nom	17.11 \ 49.80	900
Réflexion	25	42	Nom	7.89 \ 37.03	1900
Simplement	83	92	Adverbe	72.05 \ 134.46	1350
Durée	83	92	Nom	8.01 \ 19.39	400
Concentration 92)	75	100	Nom	7.53 \ 10.41	1900
Planification	83	100	Nom	0.18 \ 0.34	1900
Excellence	50	33	Nom	16.69 \ 17.43	430
Perfection	83	92	Nom	9.10 \ 16.82	1900
Réussite	25	42	Nom	11.39 \ 20.74	n/a
Démotivant	0	17	Adjectif	n/a	1400
Résultat	58	83	Nom	52.71 \ 38.38	175
Procrastination	0	17	Nom	n/a	1900

Attentes	17	25	Nom	22.53 \ 63.85	n/a
----------	----	----	-----	---------------	-----

* Ces taux de fréquence ont été tirés du site « www.lexique.org », subventionné par le CNRS. Les taux de fréquence présentés ici (TFF : taux de fréquence films; TFL : taux de fréquence livres) ont été calculés selon la formule suivante : la somme des fréquences du lemme et de ses formes fléchies sur 1 million d'occurrences. Pour de plus amples renseignements sur le calcul de ces taux, se référer au site.

** Productivité compilée par Catherine-Marie Longtin (2004).

*** D'après des analyses ANOVA, tous les mots en caractères gras sont des mots discriminant les deux groupes (la différence au niveau du pourcentage de réussite de ces mots étant statistiquement significatives).

APPENDICE B

PROTOCOLE DE RECHERCHE

PROTOCOLE DE RECHERCHE
STIMULATED RECALL EXPERIMENT ON LINGUISTIC CHOICES

Objectifs :

- a) Comprendre les choix linguistiques des participants
- b) Faire une liste des stratégies (linguistiques ou non) que les participants ont utilisées pour compléter le troisième C-Test.
- c) Vérifier si les apprenants ont utilisé les indices textuels pour faire leurs choix linguistiques.
- d) Au niveau de la morphologie : faire une liste des stratégies utilisées pour choisir le suffixe des mots dérivés.

Scénario :

A- Responsabilités de la personne qui va effectuer la recherche :

- Donner les instructions aux participants.
- Interagir avec chaque participant pendant l'activité.
- Vérifier que l'enregistreuse fonctionne bien.
- Identifier le participant par son numéro sur l'enregistreuse.
- Prendre des notes au besoin.
- Répondre à toutes les questions des participants à la fin de la rencontre.

B- Instructions pour la personne qui va effectuer la recherche :

- Mettre chaque participant à l'aise en posant quelques questions avant de commencer l'expérience.
- Présenter le déroulement de la rencontre : 3^e exercice, retour, résultats, suggestions
- Présenter le 3^e exercice et lire la consigne.
- Calculer le temps : 15 minutes
- S'assurer que chaque participant dépose son crayon dans le délai prescrit.
- Présenter le déroulement du retour + consentement de l'enregistrement
- Démarrer l'enregistreuse
- Donner les instructions pour le retour.
- Interagir avec le participant afin de favoriser la réflexion.
- Respecter les consignes du retour
- Terminer le retour en posant les questions générales
- Fermer l'enregistreuse
- Présenter les résultats des autres exercices
- Cibler les erreurs à la lumière du retour
- Apporter des suggestions quant aux choses à faire afin de s'améliorer

Instructions :

- Une fois que l'enregistreuse est ouverte, donnez la consigne au participant en ces termes :

« Nous allons maintenant regarder ensemble l'exercice que vous venez de faire. J'aimerais que vous me disiez comment vous avez fait pour compléter les mots; à quoi vous avez pensé pour trouver les mots. Vous pouvez utiliser les exemples dans le texte. »

- Assurez-vous que chaque participant a émis un commentaire sur les mots suivants :
(voir liste correspondant à chaque C-Test)

- Si le participant ne donne pas spontanément de réponse, posez-lui des questions du genre :

*« Comment avez-vous trouvé ce mot X ?
 Comment avez-vous fait ? Comment avez-vous procédé ?
 À quoi avez-vous pensé pour trouver ce mot ? »*

- Si le participant répond qu'il ne sait pas, ne pas insister. Poursuivre avec un nouveau mot ou réitérer la question initiale.
- N'essayez pas de poser des questions qui pourraient être d'un autre ordre. N'intégrez pas d'éléments de réponses dans vos questions, ni dans vos commentaires. Vos seules réponses devraient être : *« Oui / C'est bon ! / D'accord / mhm / ok / Je vois »* et vos seules formules devraient être :

*« À quoi avez-vous pensé pour trouver ce mot ? »
« Qu'avez-vous pensé pour trouver ce mot ? »
« Comment avez-vous fait pour trouver ce mot ? »*

- Une fois que vous vous êtes assurés que vous avez une réponse pour chaque mot mentionné de la liste, terminez en posant les questions générales suivantes :

*« Deux petites questions pour terminer :
 a) Qu'avez-vous pensé de ce type d'activité ?
 b) Avez-vous l'impression d'avoir appris quelque chose? »*

Répartition du temps :

1. Brise-glace :	5 minutes
2. 3 ^e C-Test :	15 minutes
3. Retour :	15 minutes
4. Résultats + erreurs :	15 minutes
5. Trucs :	<u>10 minutes</u>
Total	60 minutes

APPENDICE C

PROJET DE RECHERCHE

LA MORPHOLOGIE DÉRIVATIONNELLE CHEZ LES APPRENANTS CHINOIS ADULTES DU FRANÇAIS LANGUE SECONDE

DÉFINITION

La morphologie dérivationnelle est l'étude des mots dont la base, ou radical, a été transformée par l'ajout d'un préfixe et/ou d'un suffixe pour créer un mot nouveau.

Voici deux exemples :

Suffixation :	<i>nation</i>	+	<i>al</i>	=	<i>national</i>
Préfixation :	<i>Inter</i>	+	<i>national</i>	=	<i>international</i>

OBJECTIFS DU PROJET DE RECHERCHE

Objectif premier : Observer comment les apprenants du français langue seconde reconnaissent, manipulent et utilisent les affixes (suffixes et préfixes) pour former de nouveaux mots.

Objectif second : Identifier les principales erreurs au niveau de la morphologie dérivationnelle chez les apprenants du français langue seconde.

Objectif tierce : Identifier les principales erreurs au niveau de la morphologie dérivationnelle chez les apprenants chinois (et comparer ces erreurs avec celles produites par les apprenants d'une nationalité autre que chinoise.)

CLIENTÈLE VISÉE

Tous les apprenants adultes des cours avancés. Ces cours font partis du « Certificat en français écrit pour non-francophones » offert par l'École de langues.

MÉTHODOLOGIE

Première étape - Présentation du projet aux étudiants

Si le projet est autorisé par la responsable du programme de français de l'École de langues, la première étape consistera en la présentation du projet aux apprenants concernés. Ensuite, tous les étudiants intéressés devront signer une feuille d'entente nous autorisant à utiliser leurs tests pour les analyser et ce, bien sûr, sous le couvert de l'anonymat.

La présentation du projet, que ce soit à l'oral ou à l'écrit, se fera en plusieurs langues : français, anglais, espagnol, chinois, arabe, ukrainien, russe, etc. La durée de la présentation du projet sera d'une dizaine de minutes.

Deuxième étape - Les tests

Nous aimerions évaluer les étudiants trois fois pendant leurs études. Pour la session d'automne 2004, nous ferons passer deux tests, soit au début et à la fin de la session. Ensuite, nous évaluerons les étudiants une dernière fois à la fin de la session d'hiver 2005. Seules les copies des étudiants qui auront signé l'entente seraient récoltées.

La durée de chacun des tests sera approximativement d'une quinzaine de minutes.

RÉSULTATS

Une fois les résultats compilés et analysés, une présentation du projet sera proposée à tous les enseignants intéressés de l'École de langues. Le document écrit sera également disponible pour consultation.

AVANTAGES POUR L'ÉCOLE DE LANGUES

Premier avantage: Le C-Test est un outil d'analyse simple mais très efficace pour vérifier si un objectif de grammaire précis a été acquis (et maîtrisé). Il nous fera plaisir de donner un petit atelier pour l'expliquer aux enseignants intéressés à l'utiliser dans leurs cours.

Deuxième avantage : Les résultats de cette recherche nous fourniront énormément de renseignements sur tous les apprenants en général, mais plus précisément sur les apprenants chinois. Le résultat des analyses permettra aux enseignants de connaître les erreurs les plus fréquentes au niveau de la morphologie dérivationnelle et d'identifier les stratégies utilisées par leurs étudiants pour créer un mot « dérivé » et/ou pour éviter d'avoir à créer un mot « dérivé ».

Cette recherche se fera sous la supervision de Mme Lori Morris et Mme Marie Labelle du département de linguistique.

RÉFÉRENCES

- Altarriba, Jeannette, Judith F. Kroll, Alexandra Sholl, et Keith Rayner. 1996. « The influence of lexical and conceptual constraints on reading mixed-language sentences : Evidence from eye fixations and naming times ». *Memory & Cognition*, vol. 24, no 4, p. 477-492.
- Anckaert, P., et R. Beeckmans. 1992. « Le C-Test. Difficulté intrinsèque, pouvoir discriminant et validité de contenu ». In *Der C-Test. Theoretische Grundlagen und praktische Anwendungen*, sous la dir. de R. Grotjahn, p.145-172. Bochum : Brockmeyer.
- Bauer, Hannspeter. 1989. « MC-Test und C-Test : Die Philosophie und die Korrelation ». In *Zwischen Empirie und Machbarkeit. Erstes Symposium zum Bundeswettbewerb Fremdsprachen*, sous la dir. de Thomas Finkenstaedt et Konrad Schroder, p. 5-16. Augsburg : Universitat.
- 1992. « C-Tests auf der Volkshochschule ». *Zielsprache Englisch*, vol. 22, no 3, p. 22-23.
- Bialystok, E. 1981. « The role of conscious strategies in second language proficiency ». *Modern Language Journal*, vol. 65, no 1, p. 24-35.
- Bloom, B. 1954. « The thought processes of students in discussion ». In *Accent on teaching : Experiments in general education*, sous la dir. de S.J. French, p. 23-46. New York : Harper.
- Bourdin, Béatrice, Christel Leuwers, Carole Bourbon, Jean-Pierre Chevrot, et Marie-Anne Schelstraete. 2004. « Acquisition de l'accord en genre de l'adjectif en français écrit ». *Le Langage et l'homme*, vol. 39, no 2, p. 35-43.
- Bouthat, Chantal. 1993. *Guide de présentation des mémoires et thèses*. Montréal, Québec : Université du Québec à Montréal.
- Brousseau, Anne-Marie et Emmanuel Nikemia. 2001. *Phonologie et Morphologie du français*. Saint-Laurent, Québec : Fides.
- Brysbaert, Marc, Ilse Van Wijnendaele, et Wouter Duyck. 2002. « On the temporal delay assumption and the impact of non-linguistic context effects : Comment ». *Bilingualism: Language and Cognition*, vol. 5, no 3, p. 199-201.

- Carlisle, Joanne F. 1988. « Knowledge of derivational morphology and spelling ability in fourth, sixth, and eight graders ». *Applied Psycholinguistics*, vol. 9, no 3, p. 247-266.
- Carlisle, Joanne F. et Lauren A. Katz. 2006. « Effects of Word and Morpheme Familiarity on reading of Derived Words ». *Reading and Writing: An Interdisciplinary Journal*, vol. 19, no 7, p. 669-693.
- Carlisle, Joanne F. et D. M. Nomanbhoy. 1993. « Phonological and morphological awareness in first grade ». *Applied Psycholinguistics*, vol. 14, no 2, p. 177-195.
- Carroll, John B. 1987. « Review of Klein-Braley, C. and Raatz, U. 1985 : C-Tests in der Praxis ». In *Fremdsprachen und Hochschule, AKS Rundbrief 13/14*, Bochum : Arbeitskreis Sprachenzentren. *Language Testing*, vol. 4, no 1, p. 99-106.
- Casalis, Séverine et Marie-France Louis-Alexandre. 2000. « Morphological analysis, phonological analysis and learning to read French : a longitudinal study ». *Reading and Writing : An Interdisciplinary Journal*, vol. 12, no 3-4, p. 303-335.
- Chapelle, Carol A. 1994. « Are C-tests valid measures for L2 vocabulary research ? » *Second Language Research*, vol. 10, no 2, p. 157-187.
- Chulasai, Dusadee. 1996. « C-Test, not a Cloze Test: Test construction ». *PASAA: A journal of language teaching and learning in Thailand* [Bangkok], vol. 19, no 1, p. 53-60.
- Clark, E. V. 1993. *The Lexicon in Acquisition*. Cambridge : Cambridge University Press.
- Clarke, D. F. et I. S. P. Nation. 1980. « Guessing the meanings of words from context : Strategy and techniques ». *System*, vol. 8, no 3, p. 211-220.
- Cohen, A. 1996. « Verbal reports as a source of insights into second language learner strategies ». *Applied Language Learning*, vol. 7, no 1-2, p. 45-24.
- 1998. *Strategies in learning and using a second language*. London : Longman.
- Colé, P., C. Beauvillain, et J. Segui. 1989. « On the representation and processing of prefixed and suffixed derived words : a differential frequency effect ». *Journal of Memory and Language*, vol. 28, no 1, p. 1-13.

- Coleman, James A. 1994. « Profiling the advanced language learner : the C-Test in British further and higher education ». In *Der C-Test. Theoretische Grundlagen und praktische Anwendungen*, sous la dir. de Rudiger Grotjahn, p. 217-237. Bochum : Brockmeyer.
- Colomé, Angels. 2001. « Lexical activation in bilinguals' speech production : Language-specific or language-independent? » *Journal of Memory and Language*, vol. 45, no 4, p. 721-736.
- De Bruijn, Ellen R. A., Ton Dijkstra, Dorothee J. Chwilla, et Herbert J. Schriefers. 2001. « Language context effects on interlingual homograph recognition : Evidence from event-related potentials and response times in semantic priming ». *Bilingualism: Language and Cognition. Special Issue: The cognitive neuroscience of bilingualism*, vol. 4, no 2, p. 155-168.
- De Groot, Annette M. B. et G. L. Nas. 1991. « Lexical representation of cognates and noncognates in compound bilinguals ». *Journal of Memory and Language*, vol. 30, no 1, p. 90-123.
- De Groot, Annette M. B., Philip Delmaar, Stephen J. Lupker. 2000. « The Processing of Interlexical Homographs in Translation Recognition and Lexical Decision : Support for Non-Selective Access to Bilingual Memory ». *The Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology*, 2000, vol. 53A, no 2, p. 397-428.
- Dijkstra, Ton et Walter J. B. Van Heuven. 1998. « The BIA model and bilingual word recognition ». In *Localist connectionist approaches to human cognition. Scientific psychology series*, sous la dir. de Jonathan Grainger et Arthur Jacobs, p. 189-225. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Dijkstra, Ton et Walter J. B. Van Heuven. 2002. « The architecture of the bilingual word recognition system : From identification to decision ». *Bilingualism Language and Cognition*, vol. 5, no 3, p. 175-197.
- Dijkstra, Ton, Ellen R. A. De Bruijn, Herbert J. Schriefers, et Sjoerd Ten Brinke. 2000. « More on interlingual homograph recognition : Language intermixing versus explicitness of instruction ». *Bilingualism: Language and Cognition*, vol. 3, no 1, p. 69-78.
- Dominiek, Sandra. 1990. « On the representation and processing of compound words : Automatic access to constituent morphemes does not occur ». *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, vol. 42, no 3-A, p. 529-567.

- Dornyei, Zoltan et Lucy Katona. 1992. « Validation of the C-test amongst Hungarian EFL learners. *Language Testing*, vol. 9, no 2, p. 187-206.
- Duffield, N., L. White, J. Bruhn de Garavito, S. Montrul, et P. Prevost. 2002. « Clitic Placement in L2 French : Evidence from Sentence Matching ». *Journal of Linguistics*, vol. 38, no 3, p. 487-525.
- Eckes, Thomas, et Rudiger Grotjahn. 2006. « A Closer Look at the Construct Validity of C-Tests ». *Language Testing*, vol. 23, no 3, p. 290-325.
- Ellis, R. 1997. *Second Language Acquisition*. Oxford : Oxford University Press.
- Ericsson, K. A. et H. A. Simon. 1987. « Verbal reports on thinking ». In *Introspection in second language research*, sous la dir. de C. Faerch et G. Kasper, p. 24-53. Clevedon : Multilingual Matters.
- Faerch, Claus et Gabriele Kasper. 1980. « Processes and Strategies in Foreign Language Learning and Communication ». *Interlanguage Studies Bulletin Utrecht*, vol. 5, no 1, p. 47-118.
- 1987. *Introspection in second language research*. Clevedon, England: Multilingual Matters.
- Feldmann, Ute et Brigitte Stemmer. 1987. « Thin_____ aloud a_____ retrospective da_____ in C-te_____ taking : diffe_____ languages – diff_____ learners – sa_____ approaches ? » In *Introspection in second language research*, sous la dir. C. Faerch et G. Kasper, p. 251-267. Clevedon : Multilingual Matters.
- Frost, R., L. Katz et S. Bentin. 1987. « Strategies for visual word recognition and orthographic depth: A multilingual comparison ». *Journal of Experimental Psychology : Human Perception and Performance*, vol. 13, p. 104-115.
- Gass, Susan M. et Larry Selinker. 2001. *Second language acquisition : An introductory Course*. Mahwah, NJ : Lawrence Erlbaum Associates.
- Gass, Susan M. et A. MacKey. 2000. *Stimulated recall methodology in Second Language Research*. NJ : Lawrence Erlbaum Publishers.
- Green, David W. 1986. « Control, activation, and resource : A framework and a model for the control of speech in bilinguals ». *Brain & Language*, vol. 27, p. 210-223.

- Grotjahn, Rudiger. 1987. « How to Construct and Evaluate a C-Test : A Discussion of Some Problems and Some Statistical Analyses ». *Quantitative Linguistics*, vol. 34, p. 219-253.
- 1994. *Der C-Test. Theoretische Grundlagen und praktische Anwendungen*. Bochum : Brockmeyer.
- Gu, Y. et R. K. Johnson. 1996. « Vocabulary Learning Strategies and Language Learning Outcomes ». *Language Learning*, vol. 46, no 4, p. 643-679.
- Haastrup, Kirsten. 1987. « Using Thinking Aloud and Retrospection to Uncover Learners' Lexical Inferencing Procedures ». In *Introspection in second language research*, sous la dir. de C. Faerch et G. Kasper, p. 197-212. Clevedon, England : Multilingual matters.
- Hatta, T. 1992. « The effects of Kanji attributes on visual field differences : Examination with lexical decision, naming and semantic psychology, neuropsychology and linguistics ». *Neuropsychologia*, vol. 30, no 4, p. 361-371.
- Heilenman, L. K. et J. L. MacDonald. 1993. « Processing Strategies in L2 Learners of French : The Role of Transfer ». *Language Learning*, vol. 43, no 4, p. 507-557.
- Henderson, L. (Éd.) 1984. *Orthographics and reading : Perspective from cognitive psychology, neuropsychology and linguistics*. Hillsdale, NJ : Erlbaum.
- 1985. « Issues in the modeling of pronunciation assembly in normal reading ». In *Surface dyslexia*, sous la dir. de K. E. Patterson, J. C. Marshall et M. Coltheart, p. 459-508. Hillsdale, NJ : Erlbaum.
- Holmes, V. M., et B. Dejean De la Bâti. 1999. « Assignment of grammatical gender by native speakers and foreign learners of French ». *Applied psycholinguistics*, vol. 20, no 4, p. 479-506.
- Hoover, Micheal L. et Veena D. Dwivedi. 1998. « Syntactic processing by skilled bilinguals ». *Language Learning*, vol. 48, no 1, p. 1-29.
- House, Juliane et Shoshana Blum-Kulka (Éds.) 1986. *Interlingual and Intercultural Communication. Discourse and Cognition in Translation and Second Language Acquisition Studies*. Tubigen : Narr.
- Hudelson, Sarah. 1981. « An Introductory Examination of Children's Invented Spelling in Spanish ». *NABE: The Journal for the National Association for Bilingual Education*, vol. 6, no 2-3, p. 53-67.

- Kellerman, E. 1984. « The empirical evidence for the influence of the L1 in interlanguage ». In *Interlanguage*, sous la dir. de A. Davies, C. Crier et A. Howatt. Edinburgh : Edinburgh University Press.
- Kellerman, E. et Sharwood Smith. 1986. *Crosslinguistic Influence in Second Language Acquisition*. Oxford : Pergamon Press.
- Klein-Bradley, Christine. 1985. « A cloze-up on the C-Test : a study in the construct validation of authentic tests ». *Language Testing*, vol. 2, p. 76-104.
- Koda, Keiko. 1999. « Development of L2 Intraword Orthographic Sensitivity and Decoding Skills ». *The Modern Language Journal*, vol. 83, no 1, p. 51-64.
- Ku, Yu-Min et Richard C. Anderson. 2003. « Development of morphological awareness in Chinese and English ». *Reading and Writing : An Interdisciplinary Journal*, vol. 16, no 5, p. 399-422.
- Langer, Judith A., L. Bartolome, O. Vasquez et T. Lucas. 1990. « Meaning construction in school literacy tasks : A study of bilingual students ». *American Educational Research Journal*, vol. 27, no 3, p. 427-471.
- Laufer, B. 1997. « What's in a word that makes it hard or easy : some intralexical factors that affect the learning of words ». In *Vocabulary : Description, Acquisition and Pedagogy*, sous la dir. de N. Schmitt et M. McCarthy, p. 140-155. Cambridge : Cambridge University Press.
- Leong, C. K. 1989. « Productive knowledge of derivational rules in poor reader ». *Annals of Dyslexia*, vol. 39, p. 94-115.
- Little, David et David Singleton. 1992. « The C-Test as an elicitation instrument in second language research ». In *Der C-Test. Theoretische Grundlagen und praktische Anwendungen*, sous la dir. de R. Grotjahn, p.173-192. Bochum : Brockmeyer.
- Longtin, Catherine-Marie. 2003. « Structure morphologique de surface et transparence sémantique ». Thèse de doctorat, Paris, École des Hautes Études en Sciences Sociales.
- 2004. Liste des suffixes en français. Compilation non publiée.
- MacWhinney, B. 1987. « The competition model ». In *Mechanisms of language acquisition*, p. 249-308. Hillsdale, NJ : Lawrence Erlbaum Associates.

- Mahony, D. 1994. « Using sensitivity to word structure to explain variance in high school and college level reading ability ». *Reading and Writing : An Interdisciplinary Journal*, vol. 6, p. 19-44.
- Mahony, D., M. Singson et V. Mann. 2000. « Reading ability and sensitivity to morphological relations ». *Reading and Writing : An Interdisciplinary Journal*, vol. 12, no 3-4, p. 191-218.
- Matsumoto, K. 1993. « Verbal-report data and introspective methods in second language research : State of the art ». *RELJ Journal : A Journal of Language Teaching and Research in Southeast Asia*, vol. 24, no 1, p. 32-60.
- Meara, P. 1984. « The study of lexis in interlanguage ». In *Interlanguage*, sous la dir. de A. Davies, C. Cramer et A. Howatt, p. 225-235. Edinburgh, Scotland : Edinburgh University Press.
- 1997. « Towards a new approach to modelling vocabulary acquisition ». In *Vocabulary : Description, Acquisition and Pedagogy*, sous la dir. de N. Schmitt et M. McCarthy, p. 109-139. Cambridge : Cambridge University Press.
- Miao, Xiao-chun. 1981. « Word Order and Semantic Strategies in Chinese Sentence Comprehension ». *International Journal of Psycholinguistics*, vol. 8, no 3, p. 109-122.
- Ministère de l'Immigration et des Communautés culturelles. 2007. Statistiques sur l'immigration récente. Consulté le 5 février 2007 à l'adresse suivante : <http://www.micc.gouv.qc.ca/fr/recherches-statistiques/stats-immigration-recente.html>
- Miramontes, Ofelia B. et Nancy L. Commins. 1989. « Perceived and actual linguistic competence: A descriptive study of four low-achieving Hispanic bilingual students ». *American Educational Research Journal*, vol. 26, no 4, p. 443-472.
- Morris, Lori et Marie Labelle. 2004. « Filling the blanks: What C-test results can teach us about the French processing skills of different nonnative-speaking child populations. Communication ». *Second Language Acquisition Research Forum*. State College, Pennsylvania, 14-16 octobre 2004.
- Muljani, D., Keiko Koda, et Danny R. Moates. 1998. « The Development of Word Recognition in a Second Language ». *Applied Psycholinguistics*, vol. 19, no 1, p. 99-113.

- Nagy, William E. et R. C. Anderson. 1984. « How many words are there in printed school English ? » *Reading Research Quarterly*, vol. 19, no 3, p. 304-330.
- Nagy, William E. et Patricia A. Herman. 1987. « Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction ». In *The nature of vocabulary acquisition*, sous la dir. de Margaret G. McKeown et Mary E. Curtis, p. 19-35. Hillsdale, NJ : Lawrence Erlbaum Associates Inc.
- Nagy, William E., Georgia Earnest Garcia, Aydin Y. Durgunoglu, et Barabara Hancin-Bhatt. 1993. « Spanish-English bilingual students' use of cognates in English reading ». *Journal of Reading Behavior*, vol. 25, no 3, p. 241- 259.
- Nakagawa, Esho. 1994. « Transfer of a matching and non-matching concept between recurring and non-recurring tasks in kindergarten children ». *Japanese Journal of Educational Psychology*, vol. 42, no 4, p. 383-394.
- Nation, I.S.P. 1990. *Teaching and learning vocabulary*. New York : Newbury House.
- New, Boris et Christophe Pallier. 2001. *Lexique 3*. Subventionné par le Centre National de la Recherche Scientifique. Consulté le 11 mars 2007 à l'adresse électronique suivante : <http://www.lexique.org>
- Nippold, Marilyn A. 1998. *Later Language Development: The School-Age and Adolescent Years. Second Edition*. Austin, Texas : Pro-Ed.
- Odlin, T. 1989. *Language Transfer*. Cambridge : Cambridge University Press.
- O'Malley, J. et A. U. Chamot. 1990. *Learning Strategies in Second Language Acquisition*. Cambridge : Cambridge University Press.
- Oxford, R. L. 1990. *Language Learning Strategies : What Every Teacher Should Know*. Boston : Newbury House.
- Paprocka, Urszula P. 2002. « Acquisition du temps et de l'aspect en français L2 par des élèves polonophones débutants à intermédiaires ». *Roczniki humanistyczne*, vol. 50, no. 5, p. 133-148.
- Perdue, C. (Éd.) 1993. *Adult Language acquisition : Cross-linguistic Perspectives*. Cambridge : Cambridge University Press.

- Pexman, P. M. et S. J. Lupker. 2001. « Ambiguity and visual word recognition : Can feedback explain both homophone and polysemy effects? » *Canadian Journal of Experimental Psychology*, vol. 53, p. 323-334.
- Poulisse, Nanda, Theo Bongaerts et Eric Kellerman. 1987. « The use of retrospective verbal reports in the analysis of compensatory strategies ». In *Introspection in second language research*, sous la dir. de Claus Faerch et Gabriele Kasper, p. 213-229. Clevedon, England: Multilingual Matters.
- Prevost, Philippe. 2004. « Morphological Variation in Early Adult Sconde Language French : A Cross-Sectional Study ». *EUROSLA Yearbook*, vol. 4, p. 147-175.
- Raatz, Ulrich et Christine Klein-Bradley. 1982. « The C-test – a modification of the cloze procedure ». In *Practice and problems in language testing IV*, sous la dir. de Terry Culhane, Christine Lein-Bradley et Douglas K. Stevenson, p. 113-138. Colchester : University of Essex.
- Ravid, Dorit. 2004. « Later lexical development in Hebrew : derivational morphology revisited ». In *Language development across childhood and adolescence : Psycholinguistic and crosslinguistic perspectives*, sous la dir. de R.A. Berman, Coll. TILAR (Trends in Language Acquisition Research series), p. 53-81. Amsterdam : Benjamins.
- RENOUVO. 2005. *Le millepatte sur un nénufar : Vadémécum de l'orthographe recommandée*. Saint-Léonard, Québec : De Champlain S. F. inc.
- Rey-Debove, J. 1984. « Le domaine de la morphologie lexicale ». *Cahiers de Lexicologie*, vol. 45, no 2, p. 3-19.
- Ridley, Jennifer et David Singleton. 1995. « Strategic L2 lexical innovation : case study of a university-level ab initio learner of German ». *Second Language Research*, vol. 11, no 2, p. 137-148.
- Riegel, Martin, Jean-Christophe Pellat, et René Rioul. 1944. *Grammaire méthodique du français*. Paris, France : Presses Universitaires de France.
- Ringbom, H. 1987. *The Role of the First Language in Foreign Language Learning*. Philadelphia : Multilingual Matters Ltd.
- Rusted, J. 1988. « Othographic effects for Chinese-English bilinguals in a picture-word interference task ». *Current Psychology : Research Reviews*, vol. 7, no 3, p. 207-220.
- Sasaki, Miyuki. 1991. « A multiple-data analysis of the 3.5-year development of EFL student writers ». *Language Learning*, vol. 54, no 3, p. 525-582.

- Schmitt, Norbert. 1997. « Vocabulary learning strategies ». In *Vocabulary : Description, Acquisition and Pedagogy*, sous la dir. de N. Schmitt et McCarthy, p. 199-227. Cambridge : Cambridge University Press.
- Schmitt, Norbert et Paul Meara. 1997. « Researching vocabulary through a word knowledge framework : Word Associations and Verbal Suffixes ». *Studies in Second Language Acquisition*, vol. 19, no 1, p. 17-36.
- Schwartz, Ana et Judith F. Kroll. 2006. « Bilingual lexical activation in sentence context ». *Journal of Memory and Language*, vol. 55, no 2, p. 197-212.
- Seidenberg, Mark S. et Laura M. Gonnerman. 2000. « Explaining derivational morphology as the convergence of codes ». *Trends in Cognitive Sciences*, vol. 4, no 9, p. 353-361.
- Singleton, David. 1995. « The Bilingual Lexicon ». *Applied Linguistics*, vol. 16, no 1, p. 125-127.
- 1996. « Formal Aspects of the L2 Mental Lexicon: Some Evidence from University-Level Learners of French ». *Jyvaskyla Cross-Language Studies*, vol. 17, p. 79-85.
- Singleton, David et David Little. 1991. « The Second Language Lexicon: Some Evidence from University Level Learners of French and German ». *Second Language Research*, vol. 7, no 1, p. 61-81.
- Stemmer, Brigitte. 1991. *What's on a C-test taker's mind : Mental processes in C-test taking*. Bochum : Brockmeyer.
- 1992. « An alternative approach to C-test validation ». In *Der C-Test. Theoretische Grundlagen und praktische Anwendungen*, sous la dir. de Rudiger Grotjahn, p. 97-144. Bochum: Brockmeyer.
- Swan, Michael. 1997. « The influence of the mother tongue on second language vocabulary acquisition and use ». In *Vocabulary : Description, Acquisition and Pedagogy*, sous la dir. de N. Schmitt et M. McCarthy, p. 156-180. Cambridge : Cambridge University Press.
- Taft, Marcus. 2002. « Orthographic Processing of Polysyllabic Words by Native and Nonnative English Speakers ». *Brain and Language*, vol. 81, no 1-3, p. 532-544.
- Tiee, Henry Hung-Yeh. 1986. *A reference grammar of Chinese sentences with exercises*. Tucson : University of Arizona Press.

- Turvey, M. T., L. B. Feldman et G. Lukatela. 1984. « The Serbo-Croatian orthography constraints of the reader to a phonologically analytic strategy ». In *Orthographies and reading : Perspectives from cognitive psychology, neuropsychology, and linguistics*, sous la dir. de L. Henderson, p. 81-90. Hillsdale, NJ : Erlbaum.
- Tyler, A. et W. Nagy. 1989. « The acquisition of English derivational Morphology ». *Journal of memory and language*, vol. 28, no 6, p. 649-667.
- Tzeng, O. J. L. et W. S. -Y. Wang. 1983. « The first two R's : The way different languages reduce speech to script affects how visual information is processed in the brain ». *American Scientist*, vol. 71, no 3, p. 238-243.
- Van Hell, Janet G. 2002. « Bilingual word recognition beyond orthography : On meaning, linguistic context and individual differences ». *Bilingualism : Language and Cognition*, vol. 5, p. 209-212.
- Van Heuven, W., A. Dijkstra et J. Grainger. 1998. « Orthographic neighborhood effects in bilingual word recognition ». *Journal of Memory and Language*, vol. 39, no 3, p. 458-483.
- Vannest, Jennifer, Raymond Bertram, Juhani Jarvikivi, et Jussi Niemi. 2002. « Counterintuitive cross-linguistic differences : More morphological computation in English than in Finnish ». *Journal of Psycholinguistic Research*, vol. 31, no 2, p. 83-106.
- Von Studnitz, R. E. et David W. Green. 2002. « Interlingual homograph interference in German-English bilinguals : Its modulation and locus of control ». *Bilingualism : Language and Cognition*, vol. 5, p. 1-23.
- Wang Chunmao, Danling Peng, Linchu Guan, et Peizi Kuang. 1999. « The roles of surface frequencies, cumulative morpheme frequencies, and semantic transparencies in the processing of compound words ». *Acta Psychologica Sinica*, vol. 31, no 3, p. 257-265.
- Wang, M. et K. Koda. 2005. « Commonalities and Differences in Word Identification Skills Among Learners of English as a Second Language ». *Language Learning*, vol. 55, no 1, p. 71-98.
- Wang, M., K. Koda, et C. A. Perfetti. 2003. « Alphabetic and nonalphabetic L1 effects in English word identification : a comparison of Korean and Chinese English L2 learners ». *Cognition*, vol. 87, no 2, p. 129-149.
- Zhang, Lian Zhong. 1985. *The C-Test in Chinese*. Unpublished Ms., 26 pp.